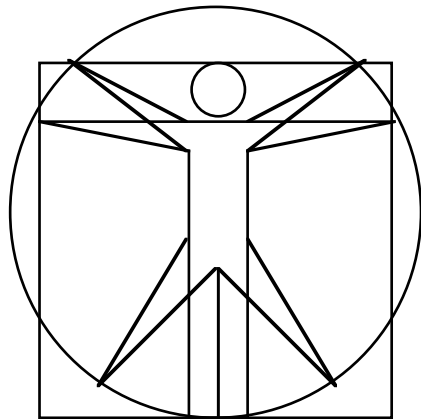


# Development of a Computerized Handbook of Architectural Plans

Ontwikkeling van een gecomputeriseerd  
handboek van architectonische  
plattegronden



Proefschrift ter verkrijging van de graad van doctor aan de  
Technische Universiteit Delft, op gezag van de Rector  
Magnificus, prof.dr.s.P.A.Schenck, in het openbaar te  
verdedigen ten overstaan van een commissie aangewezen  
door het College van Dekanen op 10 mei 1990 te 14:00 uur  
door

Alexandros Koutamanis

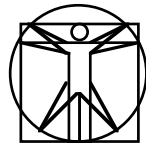
geboren te Thessaloniki, Griekenland, architect ingenieur

Dit proefschrift is goedgekeurd door de promotoren

Prof. A. Tzonis

Prof. S.J. Doorman M.Sc.

# Contents



- 1 Introduction6
  - 1.1 Computerization of visual / spatial architectural representations
  - 1.2 Computers in architecture
  - 1.3 Drawing in CAAD
    - 1.3.1 Representations and implementation mechanisms
    - 1.3.2 Architectural drawing as representation
  - 1.4 Architectural design paradigms
    - 1.4.1 The implementation mechanisms fallacy
    - 1.4.2 The holistic / multiple representation fallacy
    - 1.4.3 The sum-of-parts fallacy
  - 1.5 Why classicism
  - 1.6 An outline of the dissertation
    - 1.6.1 Chapter 1. Introduction
    - 1.6.2 Chapter 2. Recognition of locations in architectural plans
    - 1.6.3 Chapter 3. The spatial arrangement of locations in architectural plans
    - 1.6.4 Chapter 4. Cognitive filtering
    - 1.6.5 Chapter 5. An epilogue
  
- 2 Recognition of locations in architectural plans
  - 2.1 Introduction
    - 2.1.1 Spatial primitives in architectural plans
    - 2.1.2 Locations
    - 2.1.3 An outline of the chapter
  - 2.2 Input of architectural plans in machine environment
    - 2.2.1 Digitization
    - 2.2.2 Preprocessing
  - 2.3 Approaches to the recognition of locations
    - 2.3.1 Line following
    - 2.3.2 Vertices in human and computer vision
  - 2.4 Identification of the vertices of a location
    - 2.4.1 A typology of vertices
    - 2.4.2 Identification of the vertices of a location
    - 2.4.3 Why the list of vertices of a location is not directly used for the identification of the

- shape of the location
- 2.5 Recognition of shape through decomposition
  - 2.5.1 Approaches to shape decomposition
  - 2.5.2 Quadrees, medial axis transforms and generalized cones
  - 2.5.3 Decomposition into rectangular slices
  - 2.5.4 Relationships between the components of a location
  - 2.5.5 Recognition and spatial arrangement of component groups
- 2.6 Conclusions and extensions
  - 2.6.1 Beyond orthogonality
  - 2.6.2 Addressing shape variety in architectural plans
  - 2.6.3 Decomposition

### 3 The spatial arrangement of locations in architectural plans

- 3.1 Introduction
  - 3.1.1 From recognition of locations to recognition of relationships between locations
  - 3.1.2 An outline of the chapter
  - 3.1.3 Approaches to the representation of architectural plans: the structure of descriptions
- 3.2 The nature of grouping relationships in architectural plans
  - 3.2.1 An example of bilateral relationships: an adaptation of chain coding
  - 3.2.2 Multilateral interpretation of bilateral relationships
- 3.3 Types of relationships between locations
  - 3.3.1 The inheritance of Gestalt theory
  - 3.3.2 On the applicability of generalized cones and medial axis transforms
  - 3.3.3 Multilateral relationships between locations in chap
- 3.4 Recognition of grouping relationships
  - 3.4.1 The first level of group recognition: alignment, transposition and alternation
  - 3.4.2 The second level of group recognition: complex location groups
- 3.5 Preliminary conclusions

### 4 Cognitive filtering

- 4.1 Introduction
  - 4.1.1 The interpretation of descriptions

- 4.1.2 An outline of the chapter
- 4.2 Criteria of well-formedness
  - 4.2.1 The evaluation of figural goodness
  - 4.2.2 The classical canon
- 4.3 Cognitive filtering and recognition of posture
  - 4.3.1 Recognition within Marr's framework
  - 4.3.2 A multilevel system of cognitive filters
  - 4.3.3 Low level cognitive filters
  - 4.3.4 Intermediate level cognitive filters
  - 4.3.5 High level cognitive filters
  - 4.3.6 Conclusions
- 4.4 Cognitive filtering in chap
  - 4.4.1 The purpose of cognitive filtering in chap
  - 4.4.2 Cognitive filtering on the level of individual locations: proportions
  - 4.4.3 Cognitive filtering on the level of location groups:  
tripartition
  - 4.4.4 Cognitive filtering on the level of the whole architectural plan,  
I: tripartition
  - 4.4.5 Cognitive filtering on the level of the whole architectural plan,  
II: symmetry
- 4.5 Cognitive filtering: conclusions and extensions

## 5 An epilogue

- 5.1 Synopsis
- 5.2 The Intelligent Architect, the Architectural Thesaurus and chap
- 5.3 Extensions to CAAD
- 5.4 Architectural typology
- 5.5 The structure of automated design systems

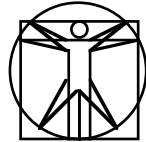
## References

## Samenvatting

## Curriculum vitae

# CHAPTER 1

## INTRODUCTION



### 1.1 Computerization of visual / spatial architectural representations

One of the most inspiring promises of current work in computer science is the integration of extensive knowledge bases into computer systems so as to create truly intelligent aids to problem solving. A prerequisite to this integration is the development of appropriate computer representations, fully compatible to the representations used in human problem solving and at the same time 'true to the nature' of digital computers.

In architecture the development of appropriate computer representations is a matter of great urgency and high priority because architecture relies heavily on complex visual / spatial representations which are rather poorly served by current computer tools. Such tools either concentrate on the lower, mainly geometrical levels at the expense of the higher ones that accommodate the principal knowledge structures conveyed by the representation or altogether ignore the issue of visual / spatial architectural representations.

The dissertation presents an approach to the development of visual / spatial computer representations for architectural purposes through the investigation of the feasibility and structure of the computerized handbook of architectural plans (*chap*), an intelligent computer system capable of recognizing the metric properties of architectural plans. The use of *chap*, an application program, in the investigation of a general methodology goes beyond the utility of an implementation as a testcase for the proposed approach. It also goes beyond that the methodology is principally directed toward the development of effective

and reliable computer tools for architecture. A more basic reason is that the coherence, reliability and accuracy required of such a fundamental tool as CHAP are helpful in ensuring that the overall methodology is characterized by the same qualities, i.e., that it is not a loose collection of random even if effective techniques. CHAP is a system for the automated recognition of visual / spatial architectural representations and hence presents a more comprehensive picture than what could be achieved by concentrating on the automated production of designs using these representations, as in the mainstream of analytical and computer studies in architecture. The use of recognition as the primary vehicle in the investigation of visual / spatial architectural representations for the computer is also beneficial for the accuracy of the representations as this is not imposed upon the data but rather built out of and through recognition.

Recognition clearly shows the peculiarities of visual / spatial architectural representations. Architectural plans in particular are strange, unconventional objects for computer vision. They are descriptions produced by a highly abstract and highly conventional representation, only distantly resembling real-world images as perceived by the eye, as recorded in photographs or even as transformed into realistic line drawings or caricatures. The analysis and computerization of this representation therefore allow direct access to elaborate and complex underlying architectural knowledge systems.

The investigation of visual / spatial architectural representations discussed in the dissertation can be summarized as an introduction of computer vision to the computerization of architectural representations. Within the framework of computer vision CHAP represents a first attempt to automate recognition of architectural representations in the most essential among architectural drawings, floor plans. CHAP accepts as input digitized images of architectural plans and recognizes their spatial articulation as configurations of parcels of space (locations) on a variety of abstraction levels. The final output of CHAP is a description of the plan in terms of the grouping formations of spaces. The output includes an analysis of the conformity of the description to spatial formal rules. In the initial version of CHAP these rules are of aesthetic (so called 'stylistic') character and are derived from the canon of classical architecture only. The description returned by CHAP is an augmentation of the description of an architectural plan as a configuration of building elements, as in most architectural computer systems, which connects directly the visual / spatial representation of architectural design with even the most abstract levels of design thinking.

The presentation of a general methodology through a

specific computer system suggests that the evaluation of the approach depends principally on the efficacy of the implementation. Indeed, a dissertation on the application of computer science to a particular problem of a discipline often constitutes a report on a concrete algorithmic system implemented as a specific computer program which attempts to resolve the problem. The evaluation of the approach and of the methodology suggested by the algorithmic system principally depends on the performance of the computer program, that is, one of its implementations.

The dissertation presents neither an approach in full nor a specific implementation of the approach but rather the *development of an approach*. The exposition focuses on the automated recognition of architectural drawings and CHAP, a specific system which covers most aspects of the recognition of architectural plans, the most basic and intricate of architectural drawings. Although the dissertation does not describe CHAP in full, it specifies in detail its overall structure and each of its modules and hence forms the first phase of the development of a fully functional prototype.

The level of specification of CHAP in the dissertation is analogous to the usual task of an architect: an architect designs a building but does not necessarily construct it. Moreover, the design of a building may specify its construction to a certain degree only. Even the most specific construction drawings and accompanying verbal documentation specify precisely the qualitative and quantitative attributes of the materials of a building and their connections but normally say very little about the temporal order of construction for every building element and between different parts of a design. Such issues are generally of no direct concern to the architect who designs the building even though they matter very much to the contractor and to the client who commissions the architect. Even the abstract scale models architects use to visualize their designs in true three dimensions are quite different from the fully functional scale models used not only in computer science but also in mechanical and electronic engineering, where mock-ups are merely indications of the appearance of a design and not of its formal structure.

Similarly, the dissertation specifies CHAP to a certain degree only: the proposed approach to the recognition of architectural plans, as well as the principles and the general structure of the techniques that comply with the approach are outlined as comprehensively as possible or necessary, but the finer details of their implementation, including the precise computational tools that may be used, are considered only in brief and in abstract. To use architectural terminology, the dissertation describes the *design* and not the *construction* of CHAP and even less the *performance* of CHAP after

construction.

The level of specificity of CHAP and of the underlying methodology in the dissertation is also related to other issues. These could be summarized as the lack of an existing overall computational framework and lack of infrastructure of partial techniques for the recognition of architectural drawings. By 'computational framework' I mean something like what Marr [1982] developed for computer vision — a general, comprehensive, well-defined and substantiated methodology for a class of related problems. By 'infrastructure of techniques' I mean specific algorithms and computer programs which attend to various aspects of the problem, in the same manner that chain coding offers a basic description of a digital curve for further analysis and recognition in image processing. Despite the considerable effort put into the computerization of architecture, techniques that could be directly integrated into CHAP are very scarce. For example, although most approaches to the automated production of architectural designs depend on descriptions of plans in terms of their interior spaces, I know of only one published technique for the automated recognition of interior spaces in architectural plans [Lawson & Riley 1982].

On the other hand, computer science offers numerous general-purpose techniques and techniques developed for applications in other areas that could be adapted for the purposes of CHAP. The dissertation aims exactly at investigating the applicability of such techniques within CHAP and also how these techniques should be correlated with constraints derived from domain knowledge.

A fundamental problem in the investigation of the applicability of these techniques is that the frequency and abundance of new techniques and related tools in computer science and the eagerness and fascination with which they are adopted in academic and professional environments alike often obscure the fact that most new techniques can be considered as variations of existing techniques or rather of the approaches and methods that underlie existing techniques.

In the following exposition of the development of CHAP, the proposed approach to the recognition of architectural plans is presented through outlines of the techniques that meet the specifications for each particular task of CHAP. The computer vision techniques used for this purpose are not necessarily the latest nor the most powerful of their kind. Instead, they are mostly widely known textbook material, even though recent advances are taken into account. Later techniques practically always offer better performance but do not offer a better picture of the effectiveness, reliability and comprehensiveness of the underlying approach. Older, more extensively analysed techniques fare better in that respect (as this is precisely the

prerequisite to their becoming textbook material) and hence are preferred as (for the moment) efficiency plays a secondary role only in CHAP. In other words, the computer vision techniques which are adapted to the recognition of architectural plans are not binding for CHAP —it is their underlying approach that matters.

A significant influence behind the decision to concentrate on the level of the approach rather than that of a specific implementation has been Marr's [1982, pp. 24–29] distinction between the three different levels at which an information processing device must be understood: the top level of the abstract computational theory of the device (the most critical for the effectiveness of the device), the middle level where the choice of representation and the algorithm of the device is made and the bottom level which covers the detailed computer architecture for the realization of the device. The dissertation considers CHAP and its underlying methodology somewhere between the top and middle level and therefore does not claim to present a complete computational theory for the recognition and visual / spatial representation of architectural plans.

The categorization of the dissertation with respect to Marr's three levels is basically determined by the principal aim of the dissertation: to investigate a methodology for the development of visual / spatial representations for the computerization of architectural practices, i.e., for the development of tools that can be used in practice. CHAP is a fundamental tool of this kind as it allows the intelligent and knowledgeable storage and recall of architectural plans in machine environment. In the tradition of artificial intelligence and knowledge based systems these representations are strongly based on cognitive and perceptual hypotheses derived from domain knowledge. The dissertation is involved in the investigation of such hypotheses only indirectly, through the representations considered and the proposed methodology for their development.

The treatment of domain knowledge in the dissertation follows a knowledge engineering viewpoint in the sense that CHAP is developed so as to accommodate a specific corpus of domain knowledge structured according to a specific domain theory but also offers the ability to substitute or integrate these with additional and/or alternative theories and corpora. Both the proposed methodology and CHAP are domain dependent in the sense that the proposed representations and recognition process are not determined by the techniques used —quite the opposite: the techniques are constrained by a fundamental approach to architectural design and architectural representations which determines the structure of CHAP and its subdivision into distinct

modules. The sequence of the different modules in CHAP is determined by a number of intermediate descriptions of an architectural plan which all reflect different levels of perception and interpretation. For example, recognition of locations in architectural plan (cfr. Chapter 2) is a prerequisite to the recognition of its spatial articulation and subdivision into parts (cfr. Chapter 3) which in turn is a prerequisite to the evaluation of its conformity to aesthetic constraints (cfr. Chapter 4). Similarly, within the module of recognition of locations, recognition of location position and shape presupposes an abstract description of the architectural plan which can be achieved by the skeletonization of the digitized image. A technique which bypasses these steps is not necessarily better suited to the purposes of CHAP because it achieves the same results with less computation. It simply denotes a different approach to that of CHAP, one that is perhaps indifferent to the role of perception in the manipulation of man-made representations such as architectural drawing.

The abstract and eclectic exposition of the computerization of visual / spatial architectural representations and their recognition in CHAP poses problems for the collection of negative (falsifying) evidence for the claims of the dissertation and also for making accurate guesses for such evidence for a future full implementation of CHAP. Even though the issue of efficiency does not enter the dissertation, an evaluation of the proposed approach could have been based on the following characteristics of CHAP and the techniques used in it:

- 1 The consistency, transparency and specificity of the representation and recognition process and the ability to integrate domain constraints (issues which relate to the reliability of the proposed approach and its modules).
- 2 The ability of the representation and the recognition process to cover all cases and forms of architectural plans that may be encountered in CHAP (issues which relate to comprehensiveness).
- 3 The support offered to the representation of architectural plans as spatial structures (issues which relate to effectiveness).

However, the lack of explicit and/or well structured knowledge bases for certain parts of CHAP (such as the grouping process of Chapter 3, especially in comparison to the cognitive filtering of Chapter 4) and the lack of direct precedents to CHAP (i.e., recognition systems for architectural representations) intensify the problems in collecting falsifying evidence because they do not allow the definition of absolute and relative measures of the reliability, comprehensiveness and effectiveness of CHAP and its

underlying methodology. To the best of my knowledge the few cases where databases of architectural plans were created in machine environment were more or less straightforward applications of CAD or pictorial database management systems with no claims on automated recognition [Brown & Steadman 1986]. As recognition is also generally ignored in computer systems for architectural purposes where all information on a design is input in an interactive manner, the comparison of CHAP and its methodology to alternative approaches is either inapplicable or unfair. Finally, the scarcity of studies on architectural perception and its relationships with formal and functional constraints makes the establishment of general criteria for the evaluation of a representation and a recognition process a rather large problem, one that is beyond the scope of the present dissertation.

As a result, any hypotheses put forward in the dissertation have a limited validity at this stage of the development of CHAP and of the underlying approach to visual / spatial architectural representations in machine environment. The reasons for that can be summarized by the main goals of the dissertation:

- 1 To investigate the feasibility of an automated recognition system for the proposed visual / spatial architectural representations.
- 2 To investigate the applicability in architecture of certain tools which have practically no opponents nor precedents among computer systems for architectural purposes and have been well established in computer vision.

In other words, the task of the dissertation in that respect is to identify fruitful directions for the development of visual / spatial architectural representations for the computer and the automated recognition of such representations, including the investigation of specific recognition techniques and descriptive formalisms. The approaches, methods and techniques that are identified in the dissertation and their investigation form the raw material for the development of precise, accurate and reliable evaluations for this new category of computerized aids to architecture. Evidence to this is that the representations and recognition process proposed in the dissertation can have beneficial effects not only to architectural practice but also to architectural theory because they offer the means to make explicit and analyse systematically the formal structures detected in or derived from all corpora of architectural knowledge (cfr. Chapter 4).

## 1.2 Computers in architecture

The digital computer offered a great challenge and high aspirations to theorists and practitioners of architecture. The late 1950's and early 1960's were marked by a popularization of computer use in academic research and in large firms. At the same time, many constituents of what is uniformly called Modernist architecture were revised through novel, more comprehensive approaches which were based on scientific disciplines and not on the fine arts, as one can see in the work of people as diverse in origin and orientation as Peter Collins and Christopher Alexander. Today the state of things is quite different, with respect to both the availability of computer resources and the general tendencies in architecture. Turnkey systems have been totally abandoned while affordable and efficient general purpose microcomputers have found their way into practically all architectural practices, while knowledge based systems are being introduced into architectural applications. At the same time, architectural theory and the concerns of the practicing architect has turned to types of investigation that are more akin to art history than design methodology.

The main question is, *What has been achieved in theoretical and practical terms since the introduction of computer-related subjects into architectural research? Can we observe patterns of progress in the utility and comprehensiveness of computer aids to architecture?* Recent accounts [Wagter 1988; Gero 1986; Gero et al 1985] have been uniformly negative with respect to the past, although they remain generally optimistic about the future, as they also were ten or more years ago. A comparison of Gero [1977] to Gero & Maher [1987] reveals goals which have changed little in the ten years that span the two overviews and dramatic changes in the means for their achievement. In addition, it seems that little has come out of the 1977 proposals and approaches that was still usable in 1987. This apparent stagnation cannot be attributed to inadequacy of means because in other fields, such as computer graphics, one can distinguish genuine progress despite a similar transition from rudimentary to quite sophisticated computer tools. Therefore, the only probable cause of stagnation is inadequate formulation of goals and of general approaches for their achievement.

The mainstream opinion is that true computer based architectural design has yet to emerge, despite some spectacular advances in computerized drawing and performance analysis. Three points are significant with respect to this opinion. First, by true computer based architectural design most if not all researchers imply *automated design systems* capable of producing either automatically or interactively complete designs from scratch. Little consideration is given to computerized design aids to

the practicing architect who is capable of producing design solutions but requires (in the sense that both needs and demands) sharper and more rational tools. A probable reason for the emphasis on complete automation is that many researchers into the computerization of architecture have been eager to prescribe reformist views of architectural design.

The second point is that what is commonly considered to be advances in computerized drawing and performance analysis is mainly merely *efficiency improvement*. In the case of computerized drawing efficiency improvements essentially rely on the advantages of the digital nature of the computer, as opposed to the analogue means traditionally employed in architecture. A digitally stored drawing can, in principle, be modified far easier than a hard copy (a drawing on paper). In the case of performance analysis, a calculation of e.g. passive solar gain can be very fast on a computer and rather slow and tedious by hand. However, it has been observed that computerized performance analysis generally reproduces rules of thumb which may be inadequate for today's complex design problems. Even worse, computer implementations of such rules fail to utilize the potential of the computer to produce and manipulate meticulous descriptions of built form, although such high degree of detail is essential in simulating the behaviour of buildings with accuracy [Maver 1988, 1986].

The third and last point is that automated production of designs in CAAD is generally based on views and approaches *imported from other fields and disciplines*. Some are based on general models of design and problems solving [e.g. Akin 1986; Akin et al 1986], while others are based on techniques developed in specific fields, in particular mathematics and linguistics [e.g. Liggett 1985; Radford et al 1985; Radford & Gero 1985; Rosenman & Gero 1985; Flemming 1987; Stiny 1980; Stiny & Mitchell 1978a]. In both cases the results have been more or less disappointing. The performance of computer based architectural design systems has never been on par with that of flesh-and-blood architects with respect to efficiency, effectiveness or reliability. Even the latest architectural expert systems offer at best coherently structured bodies of knowledge on just minor parts or aspects of design problems. The superiority of conventional architectural practices and the lack of any apparent support and advancement of these conventional practices by computer systems explain to a large degree the reluctance of practicing architects to adopt general design models proposed in the framework of research into the computerization of architecture.

Still, there are a few computer implementations of existing design techniques as well as new findings which could have been influential in architectural theory and

practice. My suggestion is that they failed to do so because they could not be integrated in the structure of domain knowledge, i.e., because their means and purposes were largely incompatible to those of conventional architectural practices. Research has concentrated on the development of design formalisms and similar reformation of fundamental and profound issues, while ignoring the necessity of in depth analysis of existing usable architectural knowledge. One example that reveals the lack of fundamental analysis of architectural knowledge is the inability of computer drawing and design systems to accommodate all stages of the design process: it is either impossible or impractical to start with vague sketches of a design solution and proceed through to the production of a final detailed description (presentation or construction drawings). Most systems require that input of design parameters is on quite detailed levels, i.e., that a large number of design decisions are taken prior to utilizing the system. Hence, they can be used only once the design has been developed to an essentially detailed level.

It can be argued that architectural research has been too hasty to expand into new ventures without first laying firm, reliable foundations on conventional time-honoured practices [Neuckermans 1987]. Intuitive elements are only too often dismissed in the computerization of architecture as subjective or even idiosyncratic and are replaced by seemingly impeccable mathematical models which may have little relevance to architecture. In related professions like law and medicine such attitudes are far less frequent with obviously beneficial results to legal and medical practice. If, for example, we compare architectural expert systems with expert systems for medical professions, the difference in performance and comprehensiveness reveal the poverty of architecture. The reasons for this poverty are not to be found with the particular expert systems but with architecture in general. Medicine relies on developed sciences which link to medical decision making through a well-founded analysis of subjective elements, such as (macroscopic) perception of symptoms by individual practitioners. Architecture, by comparison, often chooses to ignore the sciences it relies upon and forget their links with the intuitive or not apparent elements contributed by the designer. In an attempt to analyse and integrate more aspects of architectural knowledge, the dissertation follows an approach different to that of the mainstream in the computerization of architecture. A major element of this approach is the acceptance of conventional representations as a medium through which we can achieve quantification of architectural intuition.

## 1.3 Drawing in CAAD

Visual / spatial computer representations of architectural knowledge are obviously related to computer aided architectural design (CAAD). The acronyms 'CAD', 'CAAD', 'CADD', etc., often become acrostics and confuse researchers and users alike as they assume a variety of denotations and connotations in a rather wide spectrum of contexts. The basic source of the confusion is the interpretation of 'D' which might stand for 'drawing', 'draughting' or 'design'. To the user this is perhaps immaterial because in either case the final product is assumed to be primarily drawings (although it is generally accepted that computerization of drawing affects all aspects of a professional design office [Schilling 1987]). In academic circles, on the other hand, drawing is often considered as if it were a dumb, passive communication technique that can be completely separated from the intelligent process of design.

Although I do not share this opinion, it is perhaps necessary to distinguish between systems which are supposed to produce essentially drawings and those which have higher aspirations with respect to providing a structured design method. I shall term the former *drawing* systems while *CAAD* is employed as an umbrella term to denote both drawing systems and computerized design systems, as well as general descriptive techniques, such as shape grammars and rectangular arrangements. The distinction corresponds to pragmatic aspects of computerization in architecture: drawing systems are general purpose tools (the architectural profession being just a small part of their market), while *CAAD* systems are developed specifically for architects.

For the vast majority of architectural end users the only form of computerization worth investing in are drawing systems [Schilling 1987; Jones 1986]. These can produce drawings of three dimensional images far quicker and far easier than by hand, while the advantages of digital storage of two dimensional presentation and construction drawings (in particular ease of manipulation and modification) are definitely attractive. Computers are also used in project management. For such tasks the normal approach is to adapt general purpose tools to the particularities of building professions, so as to reduce paperwork through the quick and flexible production of specifications, bills of materials and other lists of more or less standardized form and nature. A most desirable feature of drawing systems is the ability to produce such lists automatically, on the basis of the database of symbols of a drawing.

The interest of professional architects in drawing is not shared by *CAAD* researchers who often stress the distinction between "dumb drawing" and "intelligent design" [Bijl

1986a, 1986b, 1982; Gero 1986; Szalapaj & Bijl 1985] or the drawing description and design knowledge [Balachandran & Gero 1988; Jain & Maher 1988] or even between “simple drawing” and “two dimensional modelling” [Port 1987], with very few dissenting voices [Neuckermans 1987; Fawcett 1986]. In fact, CAAD research has never been really concerned with computerized architectural drawing, which is always considered an annoyance, a distraction from the essence of computer aided architecture, that is, automated or semi-automated production of design solutions. The problems associated with the approaches and mechanisms underlying such aspects of automated design are discussed briefly in section 1.4. In the present section we concentrate on one particular problem of paramount importance, the *representations of architecture*.

### 1.3.1 Representations and implementation mechanisms

The issue of representation is central in chap. This is on the one hand due to the influence of David Marr's [1982] approach to computer vision and on the other relates to the nature and purpose of architectural drawing. Marr defines a *representation* as “a formal system for making explicit certain entities or types of information, together with a specification of how the system does this” [p.20], and a *description* as the result of applying the representation to a given entity. Examples of representation systems are the various numeral notations for numbers and the various alphabets for written words. The formal character of a representation should not be confused with rigidity and dissociation with the real world. As Marr points out, “To say that something is a formal scheme means only that it is a set of symbols with rules for putting them together — no more and no less” [p.21].

In the above sense, a representation is a very general and rather familiar notion that need not be based on rigorous mathematical techniques. Empirical models and everyday reference structures are equally good representations as the more accurately defined systems one encounters in e.g. syntactic (structural) pattern recognition, provided that they fulfill two essential requirements, namely that (a) resulting descriptions convey some specific aspects and properties of the described entities, and that (b) description formation is performed with consistency and efficiency. This means that a representation should be capable of describing the entities it is applied to in a manner suitable to the context of the descriptions. For example, digital computers are more

efficient when binary numerals are used, while humans prefer Arabic numerals.

In CAAD research one cannot help noticing the enormous effort put in the development of rigorous formalisms which attempt to represent the totality of design thinking in the most comprehensive manner possible. In a sense this quest is futile. To return to the last example, should we attempt to establish a unique numeral representation, common to all contexts, i.e., the human mind, the serial and the parallel computer, the calculator, the slide ruler, etc.? How could humans learn to use binary numerals efficiently? In his influential treatise on chunking Miller [1968] describes the problems computer programmers encounter in the mental manipulation of binary numerals and the transformations they employ in order to memorize lists of binary numbers. Such and similar evidence suggests that Arabic numeral representation is preferable for humans. Then, how about computers? Should computer processors use Arabic numerals, too? The answer is certainly not if efficiency of computation is of any value. Subsequently, a major problem emerges: how can we reconcile the two representations, that of the user and that of the machine? This problem has troubled computer scientists for some time. The current approach of translating Arabic numerals into binary numerals (input) and vice versa (output) through user interfaces appears to be a very sensible way of tackling the problem. It allows computer processors to operate under the binary representation, while the user deals only with familiar Arabic numerals.

In CAAD literature such cooperation of multiple representations on equal terms appears as either inappropriate to the nature of design or inefficient. Instead, holistic representations are proposed. These generally fail to recognize the partiality of a representation, that is, that each representation is effective with respect to a limited number of properties of the entities it describes. Moreover, these holistic systems underestimate drawing by considering it vague and subjective, and concentrate instead on rigid formalisms borrowed from fields as diverse as programming languages and linguistics.

We should be wary of such formalisms. Marr [1982, p.342] distinguishes between mechanisms that may be used to implement a representation and actual representations. He suggests that quite often implementation mechanisms, such as the computer programming language used in a representation, are mistaken for the representation itself. This identification is undoubtedly erroneous: the property list of a symbol in Lisp cannot be considered as a representation of the attributes of the class of entities the symbol denotes.

The distinction between representations and implementation mechanisms is a powerful criterion in the

evaluation of any proposed system. Marr applies it to, among others, Minsky's frame theory: "If frames offered a representation and not just a mechanism, we could at once see what they are capable of representing and what they are not. This may still be done, but it has not yet been; until it has, we must be wary of ideas like frames or property lists. The reason is that it's really thinking in similes rather than about the actual thing — just as thinking in terms of different parts of the Fourier spectrum is a simile in vision for thinking about descriptions of an image at different scales. It is too imprecise to be useful" [p.347].

### 1.3.2 Architectural drawing as representation

It is doubtful whether the many formalisms developed within CAAD, be they based on computer programming languages, frames and objects, or formal logic, are anything more than transfers of general purpose tools to a specific domain, i.e., implementation mechanisms devoid of any architectural knowledge. By contrast, conventional architectural drawing, although in many respects outdated, is a valid representation system for architectural design. In fact, it may be argued that it is the primary representation system for the conventional design process because it covers all formal aspects of a design solution.

The purpose of architectural drawing is to represent, not merely externalize or communicate the decisions of each individual designer, as suggested in CAAD [Lansdown 1987; Bijl 1986; Ruffle 1986]. Practically every design decision assumes its true meaning and consequences only when corresponding attributes of the design are identified in the drawings, as drawings are the documents which primarily describe a design in terms of both bulk and importance. All other documents, such as specifications and bills of materials, refer to the drawings and are incapable of describing the design by themselves.

Therefore, the distinction between "dumb drawing" and "intelligent design", as well as the insistence that CAAD research should be confined to design thinking only and exclude the final product of design, are ill-advised and unconstructive. Architectural drawing is in principle an adequate representation system for computerized design. It can cover many of the design aspects which are inaccessible to other representations, especially with respect to form — an issue totally disregarded in current architectural expert systems, where verbal description of form is often naive, inadequate or even misleading. In fact, we cannot expect that a verbal description of any kind (including the ones

structured by formalisms such as frames) could ever substitute drawings. As Black [1972, p.109] observes, “the notion of a complete verbal translation of a photograph (and still more, the notion of a verbal translation of a painting) is a chimaera.” Also in the case of architectural drawings, it seems logical to suggest (paraphrasing Black) that the information conveyed by a drawing means nothing less than what is shown by that drawing. A verbal description of the depicted entities, however meticulous and detailed, might highlight better certain aspects, make explicit certain relationships, etc., but would ultimately fail to achieve the comprehensiveness of a drawing.

The acceptance of the adequacy of architectural drawing as a representation could develop further into the statement that drawings should be the essential representation in computer aids to architecture. By ‘essential’ I mean that drawings should not merely be the output of CAAD systems but also the descriptions of the design on the basis of which decisions are taken and upon which decisions are represented (implemented) at every stage and state of the design process. Such a claim, however, could only be made on the grounds of a fully developed drawing representation for computers, such as the one proposed in the dissertation.

The issue, therefore, is how to rationalize and quantify the conventional manual architectural drawing so as to (a) use it as a representation of architectural design in an information processing environment, and (b) establish communication and compatibility with other representations that are also needed in architecture, such as those which can describe the normative thinking of architectural programmes. This calls for making explicit many aspects of drawing and above all these which have to do with abstraction and grouping. Such issues are addressed in the development of CHAP. In the present section we concentrate on a few fundamental misconceptions about architectural computer representations with respect to conventional architectural drawing.

One such misconception concerns abstraction in architectural drawing. Drawings are used at all stages and aspects of architectural design. They range from detailed and accurate construction drawings to rough, vague sketches that attempt to describe some general aspects of the solution. CAAD research underestimates the significance and expressiveness of such sketches and treats them as poor substitutes of detailed and well-founded (on comprehensive evaluations) specifications (or descriptions). Vague sketches have generally been considered as immature and incomplete versions of the final drawings. This approach reveals an extreme position with respect to architectural drawing, namely that not only is conventional design thinking hampered by the lack of analytic and other decision-

supporting tools appropriate to the complexity of current design problems, but also that more fundamental tools, such as architectural representations, are inadequate.

Although such arguments are often merely elements of propaganda for novel design approaches, there is an element of truth in them. Many of the conventional architectural design tools are outdated, especially in comparison to disciplines such as medicine and law. However, we should not be hasty in discarding existing tools in favour of novel ones; usually it is more general aspects that require reconsideration first. By adopting a more sophisticated and at the same time more pragmatic view, namely that drawing is an adequate representation to be used throughout the design process with effectiveness and reliability, our perception of architectural drawing changes. Even the vaguest and most primitive sketches become competent descriptions of the most salient features of a solution at a high degree of abstraction. This explains the reverence with which early sketches are cherished and often juxtaposed to final drawings of the same design, often with catalytic influences on our understanding of that design. Further support is provided by the fact that we *read* rather than simply *see* architectural drawings and also by the more general acceptance of the role and significance of line drawings in studies of perception: “In line drawings, the artist has not invented a completely arbitrary language: instead, he has discovered a stimulus that is equivalent in some way to the features by which the visual system normally encodes the images of objects in the visual field, and by which it guides its purposive actions” [Hochberg 1972, p.70].

Computerized drawing systems generally fail to utilize the implicit power of architectural drawing, as they usually are conceptually naive (even though technically sophisticated) transfers of manual drawing practices. For example, to facilitate manual draughting and increase efficiency, templates and adhesive sheets with standardized graphic symbols are used. Such standardization does not bias the designer’s approach to the representation of his design, except of course with respect to standardization and uniformity of parts: each drawing is a single analogue description with the same expressive power whether standardized symbols were used or not. Subsequently, it can be argued that the means of manual drawing do not affect the effectiveness of the representation.

The same cannot be said of computerized drawing where similar templates in the form of menus or “CAD libraries” of graphic symbols are used to input descriptions of building elements. As a result, computerized drawings are aggregations of discrete primitive symbols which relate to each other in a limited number of relationship types (generally spatial and part/whole relationships). Therefore,

the description of a design in a drawing system is inadequately structured and can only be considered (interpreted) as if it were analogue, thereby failing to take advantage of the potential of the digital computer.

Similar shortcomings hamper most CAAD systems, despite their emphasis on the structure and rationality of descriptions. Many CAAD systems attempt to add meaning to the graphic symbols of drawing systems by attaching to the symbols verbal and numerical descriptions of properties not described by the physical appearance and position of the symbols [e.g. Balachandran & Gero 1988; Bijl 1986; Jain & Maher 1988; Mitchell & Radford 1987; Nash 1982; Oxman & Gero 1987; Schmitt 1987; Szalapaj & Bijl 1985; Flemming 1987; Stiny 1980; Stiny & Mitchell 1978a].

In the most ambitious cases the labels attached to graphic symbols often take the form of constraints which are supposed to rule the behaviour of the graphic symbols under the appropriate queries so as to simulate the behaviour of the real-world entities they denote. For example, in a drawing the precise size and shape of a wall is normally depicted with accuracy but the additional information required for e.g. an adequate description of its thermal or acoustic behaviour is normally attached to the symbol of the wall in the form of verbal or numerical annotations. One can obviously proceed further into attaching attributes which group together symbols into descriptions of higher level entities. For example, the walls, windows, doors, floor and ceiling are the boundary of a room. By stating which room has to do with which wall, window, etc., (as in object-oriented or frame-based environments) it is possible to establish constraint networks which propagate automatically the results of the modification of one wall to another or of one room to its boundary.

Aside from the problems associated with the tremendous amount of information that has to be input and the implicit fallacy that an entity can be described as the sum of its parts (cfr. section 1.4.2), a major problem arises: how do we describe the higher level entities, i.e., those which do not correspond one-to-one to one or more specific graphic symbols in the drawing? At this point we are joined by the other major category of CAAD systems, those not concerned with drawing output at all but only with the higher (normative) levels of design thinking [e.g. Akin 1986; Akin et al 1986; Gero & Maher 1987; Gero et al 1985; Oksala 1987; Radford & Gero 1985; Radford et al 1985]. In order to describe higher level design entities, both categories inevitably revert to other representations which are considered to be better suited for the normative aspects of design thinking. This amounts to a gap between the normative levels and the final drawing product: after discussing about a design in terms of performance norms and

specifications, the user is often suddenly provided with a finished drawing without adequate explanation on how this drawing came together, on how and which parameters were deduced or translated into form [Krishnamurti 1986; Rosenman & Gero 1985; Schmitt et al 1986; Szalapaj & Bijl 1985]. Or it is required that the designer interacts continuously with the system and guides its solution with respect to form, thereby restricting the computer's role to suggesting / evaluating several options at each step and to facilitating the execution of local drawing operations [Oxman & Gero 1987].

The gap between normative levels and drawing in CAAD is to a large degree a result of underestimating the representational power of architectural drawing and in particular of its potential of abstraction. In mainstream CAAD drawings appear incapable of accommodating anything more than descriptions of a design in terms of its building components. Drawings are restricted to presenting what is considered to be verisimilar images of the proposed solution after all specifications of form have been firmly and meticulously determined. This restriction of drawing representation is first of all dangerous because it does not acknowledge the partiality of representations. By developing holistic representations for normative design thinking CAAD researchers fail to acknowledge that these representations are not necessarily appropriate for combining normative decisions into specific physical forms. Secondly, they fail to build on the body of architectural knowledge encoded in architectural drawing. If we compare architectural plans of the same design, each on a different scale, it becomes evident that each scale is concerned with different aspects of form on different levels of abstraction. On the scale of 1:200 the metric and topologic structure of a building are clearly stressed, while on 1:50 the materials and construction of each building element and component are described in detail.

In view of the abstraction of architectural drawing it is undoubtedly unconstructive to treat drawings as mere verisimilar images of the real thing. Instead, we should accept drawings as symbolic descriptions of architectural design, based on sophisticated abstraction mechanisms which can account for the different levels of resolution and the multiple points of view involved in the design of built environment. There is no reason why normative design decisions cannot be expressed by rough sketches which approximate the resulting features of the solution, nor why a designer cannot proceed from these sketches to detailed descriptions in a progression of different scales (i.e., levels of abstraction). In fact, it seems quite safe to be able to know and manipulate formal constraints at all stages of design.

However, the use of architectural drawing as the primary representation in CAAD is not as easy as it may

seem. Conventional drawings describe many aspects of a design *implicitly*. To make these aspects explicit a CAAD system needs to know how to *interpret* architectural drawings automatically, exactly as any person familiar with architectural drawing notation. A major thesis underlying chap is that shortcomings of current drawing and CAAD systems with respect to issues of representation is due to the rash and in many cases negligent adoption and adaptation of alien techniques on the basis of incomplete and unstructured domain knowledge. Therefore, it is in depth analysis of architectural knowledge that is principally needed and not credulous acceptance and import of techniques which nevertheless could have been of assistance, as we can judge from their impact on similar to architecture disciplines, such as law and medicine, where less attention is paid to creativity and innovation and tremendous effort has been invested in the analysis and codification of principles, techniques and products of conventional practice.

Perhaps the only case we have come close to accepting drawing as an essential representation in CAAD was through *shape grammars* [Gips 1975; Stiny 1975, 1980; Stiny & Mitchell 1978a, 1978b]. Shape grammars assumed a pictorial mode of architectural design and relied on the analysis of domain knowledge and structured application of the derived rules. Despite any reservations or objections one might have with respect to the analysis and formalization of domain knowledge in shape grammars, they were generally successful in the derivation of rule systems that produced consistently the entire spectrum of permissible design solutions. Unfortunately shape grammars were preoccupied with *generating design solutions*. Had they been restricted to describing architectural drawings (as did syntactic pattern recognition with essentially the same tools), we could have expected some progress in the analysis of architectural drawing as representation of design and in the analysis of the representation of architectural knowledge in general (cfr. an analogous use of chain coding in section 3.2).

However, this progress should not be overestimated. As we can see from syntactic pattern recognition, techniques similar to shape grammars have not succeeded in describing large classes of natural patterns [Nevatia 1982, p.22; Watanabe 1985, chapter 10]. The successes of such techniques are generally restricted to domains with a limited and well-defined variety of possible patterns, such as optical character recognition. Architecture is characterized by an infinite number of patterns produced through often arbitrary transformations and this explains why shape grammars are not applicable outside minute parts of the spectrum of possible architectural forms.

## 1.4 Architectural design paradigms

Even though initial attempts to automate architectural design relied predominantly on the rationalization of conventional architectural problem solving [e.g. Alexander 1964; Eastman 1975; March 1976a], perhaps out of disappointment for the many weaknesses and inconsistencies exhibited by transformations of early analyses into automated design systems, most CAAD research was quite soon misled in a vain attempt to capture the essence of architectural design in practically every formal model established in other areas.

Common in all these formal models is an often implicit belief (probably inherited from Operations Research methods and Space Allocation systems) that architectural design proceeds sequentially from domain knowledge codified in normative rules to complete and detailed specifications which in turn correspond to a limited number of architectural forms. In practical terms this meant that little analysis of domain knowledge was required besides that demanded for the calibration of model coefficients. I shall not attempt a detailed evaluation of these models. Although many interesting techniques emerged in their implementation, the significance of these models can be stated succinctly by paraphrasing David Marr: if they had discovered anything of value to architecture it would have been immediately obvious and would have irrevocably influenced the entirety of not only architectural theory but also the ways of practicing architects who, after all, are only too eager to accept and explore novel fashionable ideas.

In the early 1980's Professor Alexander Tzonis and his collaborators at the Delft University of Technology initiated a different line of investigation. The objective was once again a computer system for the automated production of architectural plans but the approach was radically different from those which had evolved in the framework of CAAD. The turning point was that the knowledge base of the system was not a set of abstract normative rules but a collection of *precedent solutions*. The system (codenamed the *Intelligent Architect*) could:

- “1. Accept and store architectural plans.
2. Identify the characteristic attributes of the above plans and organize them accordingly into a thesaurus of precedents, a memory of architectural knowledge.
3. Accept architectural programs in the form of ordinary discourses and structure them in the form of a problem to be solved.
4. Recognize similarity between the new problem to be solved and one or several precedent plan solutions contained in the thesaurus.

5. Make use of such similarity in discovering a solution to the new problem 'efficiently', inferring by analogy to precedent plans.
6. Augment its knowledge base by adding new precedents which are either found by the system or provided by the user"

[Tzonis & Scherpbier 1985, pp. 5–6].

The emphasis on concrete precedents had something of a purgative effect. Instead of searching blindly for norms, formulating arbitrary design strategies and descriptive formalisms, one had to concentrate on the systematic analysis and classification of specific bodies of precedent knowledge in the form of a number of examples of design solutions. In a sense this approach was a return to the root of the problem. Although clearly indebted to Artificial Intelligence research and knowledge based techniques, the proposed system was not a facile transfer of global problem solving methods to architecture, but rather an attempt to formalize conventional architectural problem solving through the utilization of new tools. The acknowledged affinities to the essence of the debate on typology in architecture [Tzonis & Scherpbier 1985, p.18], for instance, explain both the structure and the function of the thesaurus of precedents.

The departure from mainstream CAAD had two main constituents. Firstly, it recognized the significance of precedents as a dynamic source of domain knowledge. The acceptance of precedents as a source of architectural knowledge (a) implied freedom from rigid codifications of domain knowledge on normative levels, (b) offered the potential of informal, random augmentation of the knowledge base through the addition of more precedents, (c) supported the gradual completion of a comprehensive knowledge base (in a manner similar to the way architects are trained in architecture), and (d) ensured that the same knowledge base could be used in a variety of belief systems and from a number of perspectives through the transformation of the taxonomic structure of the collection of precedents. Secondly, the approach advocated that analysis (parsing) of precedents into structured descriptions of solution attributes was the way to recognize the parameters of a design problem. The close parallel between the parsing of precedents and the parsing of architectural programmes (normative directions) for new design problems was an implicit acknowledgement of the dangers of creating a gap between normative design thinking and drawing [Tzonis & Scherpbier 1985, p.21].

As far as I know, the closest we had come to a precedent knowledge base was probably through *rectangular arrangements* [Mitchell et al 1976; Steadman 1983]. Rectangular arrangements offered a technique for generating the whole spectrum of solution types for certain classes of problems [March 1976b, p.1]. In practical terms this meant

that for architectural plans up to a certain number of rooms there could be an exhaustive enumeration of all possible solution types. These templates could then assist the design process by e.g. providing the best matches to particular topologic requirement patterns. However, a rectangular arrangement is a *dimensionless layout* produced by a top-down technique and not a true precedent solution with precise formal and functional characteristics.

Intelligent Architect specifications called for a more pragmatic interpretation of architectural plans. Precedent plans were stored in the thesaurus module, which contained morphological characteristics of the following types:

- “1. Activity aggregation and separations in space parcels, locations.
2. Location topological relations.
  - 2.1 Graph patterns etc.
3. Location metrical relations.
  - 3.1 Grid patterns etc.
  - 3.2 Rhythmic patterns of space elements.
  - 3.3 Configurational patterns etc.”

[Tzonis & Scherprier 1985, p.23].

A more comprehensive enumeration of metric attributes listed:

- “1. Euclidean *distance* between locations. The distance from location to location considered pairwise. The information can be tabulated in the form of a matrix.
2. Relation between inside and contained outside surface giving us an overall index of *density* of a solution.
3. Coordinating metric framework: The *division* of a building into regular parts, through space constraining systems (such as grids, polar or rectilinear), defining their geometry and their limits.
4. Preordered ranking system of elements (such as the genera of classical architecture or industrial prefabricated) which specify elements according to design attributes, for example:
  - proportion,
  - contour profile, etc.
  - internal subdivision of elements
  - proportion
  - shape and profile contour
  - size
5. Relations between elements as they are placed within the metric framework constrained by:
  - rhythmic patterns of association
  - modular coordination of association
  - shape patterns of association”

[Tzonis et al 1987, p.100].

The first stages of the project were primarily focused on topologic attributes (probably under the influence of Space Allocation methods and their emphasis on aspects of circulation). Metric properties of architectural plans were reduced primarily to distance, shape and size measures [Croon 1986; Seidl 1986]. In these early attempts the role of the thesaurus of precedents was underrated. Its function was

restricted to storing and recalling descriptions of precedents on the basis of queries that required only superficial analysis of architectural plans, such as queries on aspects of size, topologic relations or pragmatic data [Seidl 1986]. The intelligent part of the Intelligent Architect system was supposed to be in the matching mechanism which related stored descriptions of precedents to sets of programmatic demands. The descriptions included in the thesaurus were merely collections of primitives representing design entities of a variety of levels and more or less commonsensical relations, such as part/whole relationships, exactly as in CAAD [Croon 1986], and thus obviously incomplete and unstructured. The lack of coherence and comprehensiveness was increased by the way descriptions were formed: a plan was described by attaching attributes to a data structure, such as filling slots in frames, either automatically (through discrete recognition systems each of which specialized on one attribute) [Seidl 1986] or even interactively [Croon 1986]. The gap between normative thinking and drawings was reemerging.

CHAP evolved out of this framework. It started as an experimental first implementation of the whole thesaurus module but soon had to retreat to an in depth analysis of just a few fundamental problems in the automated recognition of architectural plans. The purpose of CHAP is to develop a recognition process that can be completely automated and can produce comprehensive descriptions of architectural plans. The descriptions should offer multiple levels of abstraction, through which formal aspects of a design can be easily and accurately identified. In particular, CHAP is an investigation of the potential of architectural drawing as a source of information on these aspects and subsequently as a representation system for architectural form. In that respect, CHAP has to battle three major fallacies which are common to almost every attempt to computerize architecture. These are (a) the fallacy that implementation mechanisms add to the potential (or even provide the full potential) of a representation, (b) the fallacy concerning the holism / multiplicity of representations, and (c) the fallacy that a description which sums up the parts of an entity describes this entity adequately.

#### 1.4.1 The implementation mechanisms fallacy

In section 1.3.1 David Marr's approach to what constitutes representations, descriptions and implementation mechanisms was presented in brief. On the basis of this

approach I suggested in section 1.3.2 that architectural drawing is an adequate representation of architectural design. This position is not shared by the mainstream in CAAD research, who consider drawings vague and intuitive and employ instead all kinds of formalisms in order to make the representation of architectural design and analysis more rigorous than in the equivalent conventional manual / mental processes.

What is in fact implied by such formalisms is that rigorous implementation mechanisms lend their power to the actual representations. However attractive at first sight, this claim is totally unacceptable. There is no contradiction to the point that some implementation mechanisms are more appropriate than others for specific tasks because of greater flexibility, transparency, consistency or reliability. The fallacy lies in the belief that an implementation mechanism should be selected so as to compensate for the weaknesses of a representation or of a domain. We cannot expect that the inherent consistency of a design formalism derived from formal logic would improve representations of the intuitive and ill-defined problems of architectural design, nor that mechanisms imported from Chomskian linguistics would provide better access to architectural cognition because they refer to deep structures that might be common to all human mental activities. And finally, the use of frames or objects in the representation of architectural plans does not automatically ensure comprehensiveness of description, nor facilitates their classification.

Implementation mechanisms do not improve but only facilitate representation. No frame system, no shape grammar can explain why we perceive specific patterns in architectural plans, why we group two elements of a building together rather than apart, or why we fuse specific parts and aspects into higher level entities. Answers to such problems can be found only in the extensive analysis of domain knowledge and its representations. Architectural drawings admittedly offer few explicit clues for the solution of such problems but the analysis of drawing as a representation of architectural knowledge can reveal the implicit primitives and techniques of architectural design thinking.

#### 1.4.2 The holistic / multiple representation fallacy

Perhaps as a combined result of the confusion between representations and descriptions and the confusion between representation and implementation mechanisms, the general approach to representation in CAAD is Janus-like. While

there is a prevailing quest for holistic representations capable of capturing the whole essence of architectural design, in practically all CAAD systems more than one representations are used, as one can see from, among others, the gap between normative thinking and drawing.

In section 1.3.2, based on Marr's definition of representation, I argued that we should accept that in each domain more than one representations can be and are in fact used. In architectural design it is impossible to accept that analysis of programmatic demands and synthesis of form could use the same representation, unless we adopt the CAAD standpoint that architectural drawing should be irrevocably forgotten, at least in its present forms.

On the other hand, I find it equally unacceptable that within one and the same process more than one representations are simultaneously used without one being dominant over the others. In architecture drawings on different scales and of different types are used to reveal all relevant aspects of a solution, while other kinds of documents, such as specifications and bills of materials are only complementary to the drawings and generally destined to assist ancillary tasks, such as purchase and cost estimation of materials. Such documents are inadequate as descriptions of a design even for slightly more complicated tasks, such as the estimation of the duration and total cost of each aspect of construction, as these require detailed knowledge of the form of the solution. We could make even finer distinctions, such as between architectural plans and topologic diagrams. A topologic diagram corresponds to one or more plans and assists in the representation of specific functions, usually of a dynamic nature, such as circulation of people or goods. For the trained eye and mind it is not difficult to detect topologic relationships in ordinary architectural plans and thus render the use of topologic diagrams in the design process redundant.

Therefore, it is only logical to assume that, if we accept the validity of architectural drawing as a representation of architectural design, we could isolate drawings as the primary (or only) source of information about the form of a design and subsequently use them as the essential sources for the recognition of the primitives and the structure of design representation.

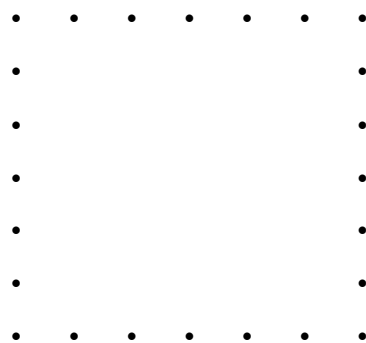
### 1.4.3 The sum-of-parts fallacy

By imposing constraints and attaching attributes to symbolizations of primitive parts of the description of a building, or by defining hierarchies based on variations of the part/whole relationship, many CAAD systems attempt to

describe design entities of a specific level of abstraction in relation to entities of other levels, and thus simulate the structure and behaviour of real-world objects. This approach largely follows a commonsensical view of the structure of the real world, namely that each entity can be described in terms of its parts. A human, for example, is a being with two legs, two arms, a torso, a head, etc.; a hand is the part of a human that is found at the end of the arm, and so forth.

From some points of view such descriptions can serve a purpose but from others they fail because transition from one level of abstraction to another usually requires restructuring of descriptions. Or, as Gestalt theory put it, the sum of parts does not equal the whole: “It has long seemed obvious ... that ‘science’ means breaking up complexes into their component elements. Isolate elements, discover their laws, then reassemble them, and the problem is solved. All wholes are reduced to pieces and piecewise relations between pieces. ... [However]: There are wholes, the behaviour of which is not determined by that of their individual elements, but where the part-processes are themselves determined by the intrinsic nature of the whole” [Wertheimer 1938, p.2].

Indeed, when we see the image



we normally perceive a rectangle (output) whose edges comprise regularly spaced bullets (input). Any attempt to describe the image in terms of just its discrete parts, the bullets, and their relative and absolute position should go to great lengths to identify the image as a rectangular four sided plane figure. First, one has to recognize the four edges in the figure and then check the relations between them with e.g. the Euclidean definition of a rectangle. This implies first of all a transformation of the bullets into edges. Grouping the bullets into four distinct edges seems at first easy but actually involves the development of a relatively large and complex rule system for the size of the task. For example, do the bullets at the four corners belong to both edges which meet there, to just one of the two or to neither edge? Any decision with respect to this implies an analogous modification of the definition of the rectangle because in Euclidean geometry

each pair of neighbouring edges in a rectangle meet (intersect) at a dimensionless point. In other words, it is necessary to consider the implications of a decision on higher levels of abstraction.

In visual / spatial architectural representations one encounters similar problems: in a sense a room is defined by its walls, openings, floor and ceiling, i.e., its boundary, and the objects it contains, such as fixed and mobile furniture. However, a description of the room in terms of just these elements is incomplete, as it would fail to convey any of the properties of the room but its physical boundary and content. Equally limited are descriptions of a room in terms of just higher level attributes, such as its topologic relations to other rooms, its performance in relation to criteria of environmental quality, etc. If both kinds of description are simultaneously used we can expect nothing more than two descriptions which, although of the same entity, have little in common. We can neither expect that the mechanisms used for the implementation of such representations provide the links between different parts and levels of description nor that there are or that there may be found better, more accurate ways of summing up parts to obtain the whole. On each level of abstraction one can perceive aspects which are not apparent in descriptions on other levels. To reveal such aspects one needs to know the fundamental cognitive transformation mechanisms used in architectural representations, such as those which abstract one level of description into another, e.g. a description of the construction of each building element (as in a drawing on 1:20) to a description of the metric and topologic structure of the whole building (as in a drawing on 1:200).

These mechanisms cannot be found outside architecture, even though assistance from other fields such as cognitive psychology and computer vision is required. The significance of architectural knowledge becomes even more evident when we consider that in an intelligent computer environment besides to abstraction we also need criteria for judging the acceptability of resulting descriptions with respect to technical or aesthetic criteria.

The sum-of-parts fallacy in CAAD assumes a rather particular denotation, as revealed by the use of terms such as "CAD-like models" in other fields [Pentland 1986, p.243]. In most CAAD systems an object is described simply in terms of collections of low-level features, such as a wireframe model of edges. The direct and generally acceptable character of such descriptions accounts for their popularity in computerized drawing systems in architecture and in other engineering disciplines and even for their introduction in computer vision (mainly in industrial applications). Despite the successes in carefully controlled and restricted environments, it is generally accepted that such

impoverished representations undermine the flexibility and reliability of the recognition process [Pentland 1986, p.243]. Similarly, we cannot expect that reliance on just low level features can form a sound basis for intelligent computer aids to architecture.

## 1.5 Why classicism

The dissertation investigates visual / spatial architectural representations through the development of a computer system for the automated recognition of metric descriptions of architectural plans. The restriction to metric properties only is merely a device to scale down the problem without reducing the applicability of the proposed approach and techniques. Descriptions from other point of views, such as topologic or functional descriptions, are either derived from or represented on the basis of metric descriptions.

Another scaling down device is the inclusion of only totally orthogonal plans in CHAP. Although a full justification of this should be based on a detailed and comprehensive model of transformations, it can be argued that differences due to orthogonality, obliqueness or curvilinearity are often only superficial. This is evident when we consider topologic aspects or systematic transformations of essentially orthogonal forms, as in the work of Alvar Aalto. The inclusion of architectural plans with oblique or curvilinear elements in CHAP would simply increase the number of primitives of the representation without altering the structure of descriptions or of associated procedures.

The most significant scaling down device is the inclusion of only classical architectural plans in CHAP. This restriction serves several purposes. Firstly, by concentrating on just one formal system it is possible to analyse in depth its fundamental characteristics and ignore the 'archaeological' approach encountered in many historical and methodological treatises. Such analysis promotes our understanding of what constitutes a formal system and a more coherent description of the characteristics of classicism.

Secondly, by concentrating on classicism it is possible to exploit the extensive analysis of classicism as found in a remarkable amount of influential treatises. Classicism is one of the few areas in architecture that has been soberly and scholarly studied practically throughout modern history, thus supplying an almost exhaustive variety of (often complementary) interpretations from a wide spectrum of viewpoints. One of these, the first implementation of shape grammars in architecture [Stiny & Mitchell 1978a], is an obvious challenge for CHAP, as it proposed both a

representation for a class of classical architectural plans and the limits of the subdomains which can be defined by the primitives of each implementation (variation) of this representation.

Thirdly, the analyses of classicism lead to the identification of a self-contained rule-based canon which constrains the production of classical architectural works. Although the dissertation is not concerned with the nature and structure of the classical canon, its rule-based character is one of the primary advantages of classicism as the domain of the initial version of CHAP because it facilitates the decomposition of the canon into a hierarchy of analyses that measure the acceptability of a classical design (cfr. Chapter 4). This decomposition is essential for (a) the integration of the classical canon in the bottom-up description and recognition of architectural plans in CHAP and for (b) the investigation of the cognitively oriented interpretation of the classical canon as a top-down rule-based generative system in Tzonis & Lefaivre [1986], CHAP's essential guide to classical architecture.

Finally, classical architectural plans are generally compatible with the other scaling-down devices, consideration of metric properties of architectural plans only and orthogonality. The classical canon is concerned primarily with form and, although it applies to the three dimensions of a building, is clearly and comprehensively expressed in floor plans. Classical plans are not always orthogonal but orthogonality is their most 'normal' state. Also, they are characterized by properties which make them attractive guinea pigs for CHAP. One of these, the clarity of the boundaries of interior spaces, is particularly advantageous for scaling down the process described in Chapter 2.

It should be stressed again that the approach to domain knowledge in the dissertation is from a knowledge engineering viewpoint, in the sense that domain knowledge—in this case, analyses and formalizations of the classical canon—are considered as resources for the investigation of representations rather than objects of investigation. The dissertation is concerned with the nature and structure of such formalizations only with respect to their integration into computer representations of architectural knowledge and into the recognition of such representations. One consequence of this is the adoption of a bottom-up approach (cfr. Figure 1-1) to the top-down structure of the classical canon in Tzonis & Lefaivre [1986]: by assuming an opposite direction to that of the proposed process of generation of classical patterns the dissertation simply aims at investigating the consequences of the generative process from another perspective, that of recognition and representation.

## 1.6 An outline of the dissertation

The sequence of chapters in the dissertation reflects the sequence of descriptive and recognition modules in CHQP (Figure 1-1) and in general the structure of the proposed recognition process of visual / spatial architectural representations.

As mentioned in the previous section (1.5) the essential source of domain knowledge [Tzonis & Lefaivre 1986] considers the classical canon as a top down generative system. Therefore, it could be argued that the rules that apply to the form of architectural plans contained in CHQP (i.e., the relevant parts of the classical canon) should be placed in the front end of the recognition system, that is, in the beginning of the derivation of the description of an architectural plan.

However, a top-down approach to recognition, besides the determinism and unreliability it entails, represents little more than a straightforward application of the model of domain knowledge it utilizes. The present dissertation is primarily an *investigation* and as such it should provide either the most extensive or the most extreme tests to the issues it considers.

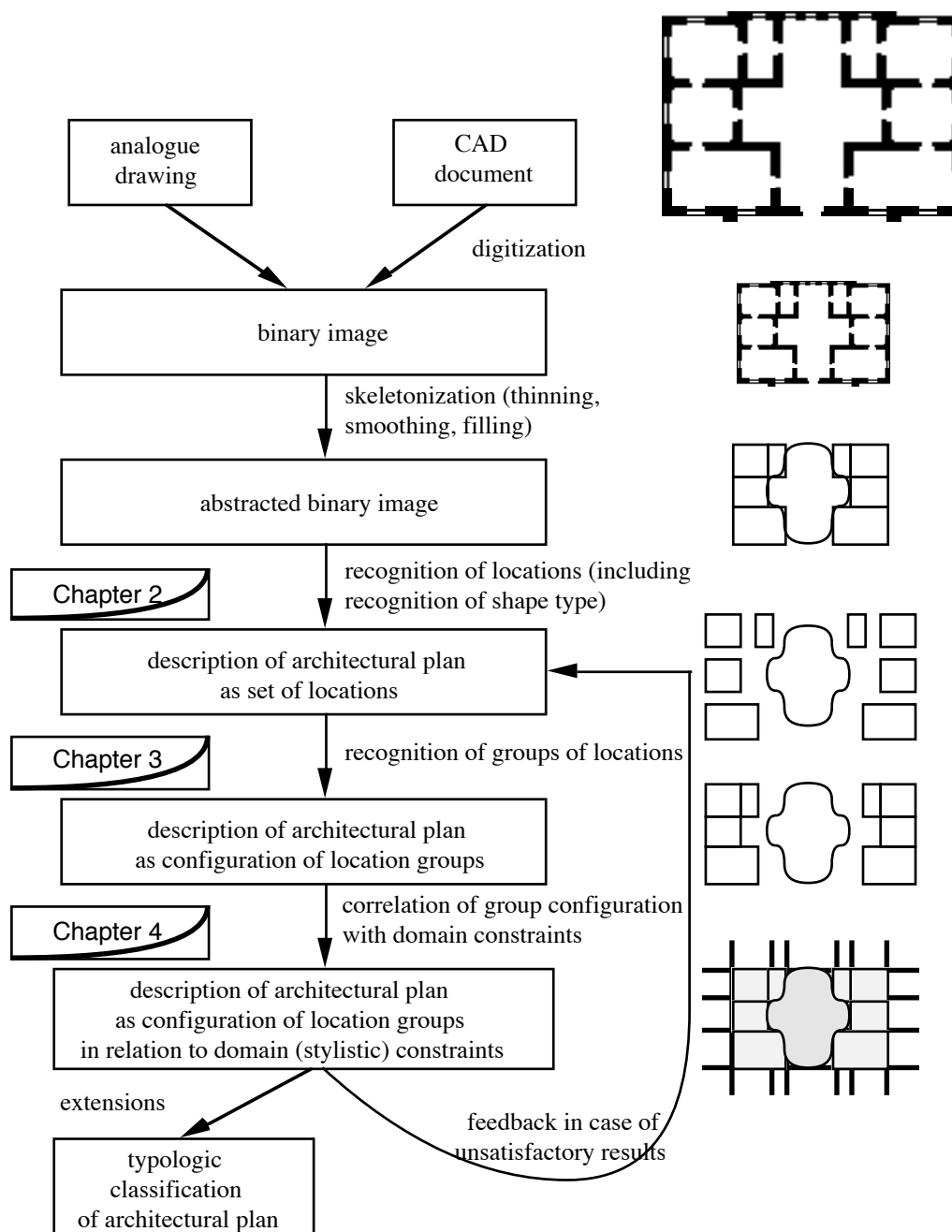
As a result, it was preferred to adopt the standard bottom-up approach of computer vision which also corresponds to our perception of architectural plans: even with a very limited knowledge of classical architecture and of the classical canon one is capable of reading and understanding an architectural plan. Knowledge of the classical canon *enhances* understanding —it does not make reading the plan possible.

### 1.6.1 Chapter 1. Introduction

Chapter 1 suggests that the poor performance of computerized architectural drawing and design systems is among others evidence of the necessity to computerize visual / spatial architectural representations and proposes that a recognition system such as CHQP offers perhaps the most comprehensive means for the investigation of a methodology for the development and use of such representations. In contrast to the mainstream in CAAD research where drawing is treated as a poor substitute for impeccable formalisms for the representation of design thinking and its products,

Chapter 1 proposes that conventional architectural drawing is an essential source of visual / spatial architectural representations for the computer.

A measure of the value of such representations and comprehensive recognition systems is their significance for the utilization of precedents in computerized architectural design. Precedents allow the comprehensive and structured organization of architectural knowledge in a manner that is closely related to human problem solving in architecture and therefore allow a modular computerization of architectural design and a direct integration of resulting computer aids into the conventional design process. CHAP was precisely developed with the requirements of a computer database of precedents in mind.



**Figure 1-1** A schematic outline of chap

Finally, Chapter 1 presents the basic scaling-down devices of the dissertation and chap:

- 1 Consideration of metric characteristics of architectural plans only.
- 2 Inclusion of only totally orthogonal architectural plans in chap.
- 3 Inclusion of classical plans only.

Of these, (1) and (2) mean that the content of the initial version of chap represents some kind of least common denominator for the whole spectrum of architectural forms and descriptions: all descriptions of all aspects of an architectural plan are derived from or refer to the description of its metric characteristics, while all oblique or curvilinear forms can be transformed into orthogonal ones through, for instance, grid substitution. In other words, (1) and (2) restrict the content of chap without reducing the scope of the underlying methodology as they simply lower the number of patterns that may be encountered without artificially eliminating any of the classes of patterns or problems that should be confronted.

On the other hand, (3) is a scaling-down device that restricts the scope of chap to architectural plans of only one formal system ('style') — a necessary restriction since each system is characterized by its own formal rules. Still, this restriction does not reduce the overall validity of the proposed approach nor the applicability of chap and its techniques and methodology to other systems. Extensions to other formal systems require little besides an augmentation of the grouping and cognitive filtering rules of the initial version of chap (cfr. Chapters 3 and 4). The structure of the initial version is the essential guide for the dissemination of domain knowledge and the derivation of the additional rules for such augmentation. Therefore, if the initial version of chap is successful in recognizing classical architectural plans, there seems to be no fundamental objection to widening its scope so as to include the whole of architecture.

## 1.6.2 Chapter 2. Recognition of locations in architectural plans

Chapter 2 describes a fundamental task of chap: the recognition of the position and shape of the locations, the atomic parts of the description of an architectural plan in chap. This operation represents the final and most significant part of the first stage in processing an image input

in machine environment.

This stage also covers the skeletonization of the digitized binary image into a skeleton, an operation which is not considered in detail because it is to a large degree implementation dependent. Chapter 2 concentrates on the transformation of an analogue description (the binary image) into a symbolic one (a list of locations), a transformation that provides the essential building material for the description and recognition of the depicted plan.

The term *locations* is used instead of the more usual (*interior*) spaces because, although within the metric concerns of chap both denote the same class of spatial primitives, locations have functional connotations which remind us that it is not only form that influences our perception and definition of atomic parcels of space.

The fundamental character of the process described in Chapter 2 signifies its critical importance for the accuracy of the description and recognition of architectural plans in chap. As a result, the primary criteria for the selection of a particular technique for the recognition of locations are reliability and comprehensiveness. These disqualify simple techniques such as line following and lead instead to the adoption of a technique that relies on the identification of the vertices of a location and propagation of the constraints of their type.

This technique relies upon the typology of location vertices and their connectivity. In a totally orthogonal environment there are only eight possible types of vertices (Figure 2-10). A vertex of one of these types may be connected in the horizontal and vertical direction to vertices of specific types only (Figure 2-11). Therefore, once one vertex of a location is identified, the expectations as to its neighbours in the horizontal and vertical direction can be formed with more speed, accuracy and flexibility than by line following.

The proposed technique relates to the perception of illusory contour figures and to the recognition of line drawings of three-dimensional scenes of trihedral objects as ultimately formalized by Waltz [1975]. And as these two, it can be used not only for totally orthogonal but for all kinds of architectural plans.

Identification of the vertices of a location is followed by recognition of the shape of the location. Within the context of the relatively simple architectural plans of the initial version of chap, the number of possible location shapes is rather limited (cfr. Stiny & Mitchell [1978a]). Therefore, the use of templates for each shape type would be a realistic solution for the initial version. However, the extension to more complicated totally orthogonal architectural plans would have required a disproportionately

large number of templates which would make the rule base of shape recognition too awkward to use and manage.

An alternative to template matching is decomposition of each location into simple rectangles and recognition of shape on the basis of the spatial relationships between these rectangles. This approach relates to similar decomposition of images for e.g. the identification of the generalized cones it consists of. The essential differences are that:

- 1 `chap` detects the relationships between the components of a form on the boundary of the form and not along medial axes.
- 2 Shape type recognition is more dependent upon the uniformity and repetition of spatial relationships between components of a location than on the number of these components.

Of these, (1) relates to the significance of boundaries to space and in particular to locations and architectural plans (as opposed to the significance of the skeletal system in animal forms for which medial axes are preferable). The recursive or iterative character of complex location shapes (as opposed to the strictly limited number of components in animal forms) makes (2) a necessary choice for the abstraction and subsequent classification of shapes on the basis of the relationships between their components.

The proposed approach to shape recognition relates to the recognition of location grouping (Chapter 3) as both rely on the same basic grouping relationships. For this reason most comparisons to precedent descriptive techniques and approaches are presented in Chapter 2 rather than in Chapter 3.

### 1.6.3 Chapter 3. The spatial arrangement of locations in architectural plans

Chapter 3 moves to the next significant problem in `chap`: recognition of spatial arrangement of locations in architectural plans, that is, recognition of grouping relationships that determine the subdivision of a plan into parts. In the absence of systematic and exhaustive typologic studies of classical architecture that would allow us to define a repertory of the location group types possible in classical architectural plans, Chapter 3 follows a bottom-up approach based on grouping relationships derived from elementary architectural knowledge and formalized with assistance from Gestalt theory and its antecedents.

The grouping process described in Chapter 3 corresponds both in purpose and in structure to the derivation of a description of an image in computer vision [Marr 1982]. This correspondence suggests that the primary criterion of the effectiveness of this process is whether it recovers the 'natural parts' of an architectural plan, that is, whether the

configuration of location groups returned by the grouping process corresponds to one of our intuitive interpretations of the subdivision of the plan.

The bottom-up approach assumed in Chapter 3 is both an alternative to the top-down generative process suggested by Tzonis & Lefaivre [1986] and an investigation of the structure and of the rules of this process. As an alternative, it suggests that perception follows a different direction to the top-down generation of classical patterns. However, as there is no concrete evidence of the influence of perception on design, the dissertation makes no specific claims with respect to the nature of the generative process of classical architecture (or any other style). Instead, the bottom-up grouping process of Chapter 3 is seen as an investigation of the consequences of the top-down generative process of Tzonis & Lefaivre, that is, as a way to verify the results of the top-down generative process.

Recognition of location groups is subdivided into two levels. The first is recognition of simple grouping relationships, namely alignment, alternation and transposition along a single direction. The resulting groups form the building blocks for the second level which investigates relationships of closure, continuation and proximity between them. On the basis of these relationships simple location groups are merged into complex groups which correspond to concrete categories of spatial arrangement in architectural plans.

The final output of the grouping process described in Chapter 3 is a configuration of location groups which describes comprehensively and coherently the architectural plan. This description is equivalent to the ones returned by most computer vision systems whereby an image is described in terms of its regions and the spatial relationships between regions. The usual approach to recognition in computer vision is to match the description to a database of known models. In *CHQP*, the absence of systematic and exhaustive typologic studies on classical architectural plans does not allow the creation of a database of actual plan types. The formalization of the classical canon by Tzonis & Lefaivre [1986] allows the creation of a comprehensive typology of classical architectural plans but only on a rather abstract level. The correlation of the abstract patterns produced by the classical canon on this level with actual plan types would still require a systematic typologic study of a rather large number of classical plans. As such a study is beyond the scope and limitations of the dissertation, correlation of the output of the grouping process with constraints of the classical canon is attempted in Chapter 4 in an analytical manner.

#### 1.6.4 Chapter 4. Cognitive filtering

The grouping process of Chapter 3 contains little explicit reference to the classical canon even though the canon constrains heavily the generation and the perception of a classical plan. Moreover, this grouping process may (and usually does) return two or more alternative descriptions for the same architectural plan. Chapter 4 addresses the problem of establishing a preference order on the basis of criteria of well-formedness which represent constraints of the classical canon.

The investigation of the well-formedness of a pattern is closely related to the issue of figural goodness and thus to notions of economy in encoding. These notions offer quite useful insights to the character of well-formedness but few clues for an explicit integration of the classical canon into them or vice versa. Chapter 4 investigates the well-formedness of the description of a classical architectural plan in a purely analytical manner. This means that each relevant level (or sublevel) of the classical canon according to Tzonis & Lefaivre [1986] is transformed into a single group of criteria of well-formedness which is investigated independently. The hierarchical structure of the classical canon determines the coordination of these criteria into a sequence of cognitive filters which progressively analyses the correspondence of the descriptions derived by the grouping process of Chapter 3 to the constraints of the canon.

Cognitive filtering starts with tripartition, the most essential of relevant levels of the classical canon. Tripartition is investigated first within each of the location groups in the description of the architectural plan. The resulting local tripartition frames are then collated into a global tripartition frame which represents the application of tripartition to the whole plan.

The global tripartition frame forms the basis for the investigation of symmetry. This investigation is conducted in a way derived from the measurement of *transformational economy* in a pattern [Garner 1974; Palmer 1982, 1983]: the symmetry of a plan is measured by its invariance with respect to the global tripartition frame under a number of geometric transformations (rotations and reflections). Invariance is measured with respect to both the number and the structure of location groups within each of the subdivisions defined by the global tripartition frame.

The output of the process described in Chapter 4 is a preference order for the descriptions derived by the process of Chapter 3 on the basis of an analysis that also allows and facilitates formal evaluation and classification of the

described architectural plan. Subsequently, the main criteria for the technique used to implement this process is the correspondence of its criteria with the meaning, nature and scope of the same or equivalent criteria in intuitive interpretations of classical architectural works.

An evaluation of the technique of cognitive filtering proposed in Chapter 4 can also be based on its application to an equivalent problem, recognition of posture (section 4.3). This application attempts to show that

- 1 Precise and detailed recognition requires more than direct matching to a database of known models. For such tasks, analysis (and possibly decomposition) of the models is required. Similarly in CHQP recognition of the type of an architectural plan could have been based on a typology of known plans. However, even if the typology was derived with direct reference to the constraints of the classical canon, recognition of type in this manner still would not allow a direct and detailed evaluation of the well-formedness of the plan.
- 2 Decomposition of the models of a model base and their reorganization into a comprehensive hierarchy of parts structured according to the relationships between these parts may allow a smoother transition from the bottom-up derivation of a description to the top-down recognition by matching to known models.

### 1.6.5 Chapter 5. An epilogue

The methodology and techniques investigated in the dissertation are primarily considered with respect to CHQP, a specific automated recognition system. The resulting specification of the structure of CHQP gives a measure of the use of such a system within the context of a computerized collection of architectural precedents and also presents several extensions to other areas of architecture. Although these extensions are not considered as verifiable claims, Chapter 5 describes some of their implications.

One implication concerns the role of architectural drawing in computerized drawing and design systems. The dissertation suggests that CAAD systems should employ automated recognition modules which would reduce the information required to be input by the user and allow more flexible and more effective communication with him.

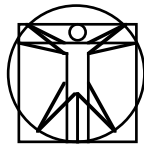
A related issue is that of architectural typologies. Although CHQP is too closely related to a specific model of domain knowledge to make any wider claims on the nature of architectural typologies, its structure suggests a bottom-up combinatorial derivation and classification of types of

architectural plans, that is, a description of each type on the basis of a limited repertory of component types (i.e., group types) and a limited number of coordinating devices (i.e., cognitive filters).

Another extension of the dissertation and chap concerns the nature and structure of systems for the automated production of architectural designs. The distinction between the derivation of a description and its recognition with respect to the constraints of classical architecture in chap or a database of known models in computer vision invites an investigation into the feasibility and performance of an analogous distinction in automated design systems. This distinction should probably go beyond the distinction between the production of a complete design and its aesthetic evaluation, as in shape grammars [Stiny & Mitchell 1978a versus Stiny & Gips 1978b].

# CHAPTER 2

## RECOGNITION OF LOCATIONS IN ARCHITECTURAL PLANS



### 2.1 Introduction

This chapter presents an approach to the recognition of the descriptive primitives used in chap. Definition of these primitives represents the first and perhaps most crucial step in the development of visual / spatial computer representations for architectural purposes because it determines the lowest level of abstraction for a representation. The approach adopted with respect to the recognition of these primitives is equally significant for the use of the representation because it signifies the process through which a description is derived and structured.

In defining the primitives for visual / spatial architectural representations the dissertation diverges from CAAD and the methodologies of computerized geometric modelling that underlie definition of descriptive primitives in CAAD. To use a couple of terms that used to be very fashionable in architecture, geometric modelling is concerned with the solid parts of an object, which in a building correspond to its constructed parts, that is, its building elements. The dissertation addresses a different level of abstraction, that of the voids of a building, that is, of space bounded by building elements. The level of 'voids' is complementary to that of 'solids': while the latter describes construction, the former is related to more abstract aspects of architectural design, such as the formal and functional articulation of a building.

The significance of primitives for a representation in general and for chap and the

representation it employs in particular leads to the secondary purpose of this chapter: a review of techniques and approaches to representation relevant to both to the recognition of primitives in CHAP, discussed in this chapter, and to their grouping, discussed in Chapter 3. Although definition and recognition of primitives is only a part of these representations and descriptive techniques, the bulk of the review is incorporated in this chapter in order to provide a more comprehensive picture of the context of each kind of primitives.

### 2.1.1 Spatial primitives in architectural plans

Architectural plans, either manually drawn or computer generated, describe a design essentially in terms of its building elements or, rather, the dimensions and position of each element and aspects particular to each element type, such as the swing of a door. These descriptions suffice for many tasks, such as production of bills of materials, but require interpretation on higher levels of abstraction in order to cover aspects other than those constrained by construction only, such as pedestrian circulation or transportation of goods.

For most such aspects the description of a design must be on the level of its spatial articulation and subdivision. For example, any case of transportation or circulation is from where one activity takes place to where another activity takes place through a sequence of other spaces. Definition and recognition of the primitives which describe these places and their relationships is a prerequisite to any description and manipulation of an architectural plan as a meaningful spatial structure rather than as a set of lines and shapes or as a set of building elements.

Descriptions of buildings as configurations of spatial primitives have been widely used in CAAD, from space allocation methods [Eastman 1975] to shape grammars [Stiny & Mitchell 1978a]. In several CAAD approaches an architectural plan is described primarily in terms of its interior spaces which are the output of all major formative operations. Operations on the level of building elements elaborate the resulting spatial configuration. The general acceptance of the predominance of descriptions on the level of spaces in automated design systems was not accompanied by the development of techniques for the automated recognition of such descriptions. Instead, recognition of spatial primitives in CAAD is normally performed in an interactive manner.

In one category of computer systems (mainly drawing

systems) the description of a design is practically the same as in manual drawing: a collection of pixels and/or vector graphic symbols which must be interpreted by the user, exactly as an analogue drawing. To make spatial primitives explicit in this manner the user must add a second description on top of the first one, for instance by describing all rooms in an architectural plan one by one, each as a two-dimensional shape which he draws on a layer other than those that contain building elements. Then the computer system can answer queries about the position, shape and size of each room and, provided that detailed connections exist between the different layers, infer certain properties of rooms. For example, it is possible to determine whether two contiguous rooms are directly linked with respect to pedestrian circulation by searching for doors among the building elements that separate them.

More advanced CAAD systems attempt to make the description of a design more comprehensive and meaningful through descriptive formalisms which integrate all relevant aspects. The justification of such formalisms is that the explicitness of multiple interrelated descriptions of the same objects can be helpful in design because it establishes an explicit structure for descriptions of design procedures and products [Jones 1970; Radford & Stevens 1987]. The major problem with computer systems that follow this approach is that little is inferred by the computer. Here again the user inputs all aspects of the description of a design albeit in a more structured manner that facilitates information processing. For example, the user inputs a complete description of each building element, from its size and position to its material and colour, as well as its connections to other building elements. These connections normally include implicit descriptions of spatial primitives: one of the properties of a wall is the room(s) to whose boundary it belongs. The user also inputs explicit descriptions of the spatial articulation and subdivision of the building which in turn include implicit descriptions of building elements: one of the properties of a room is the walls that form its boundary. The overall structure of the descriptions facilitates correlation of different aspects (in our example construction and spatial articulation) and transition from one aspect to the other but still necessitates interactive definition of all descriptions, at least on the level of the primitives each description employs.

The main disadvantage of interactive definition of all primitives is the heavy workload and burden placed on the shoulders of the user, who has to provide explicit descriptions of too many aspects of a design, while the description of a single aspect may implicitly contain descriptions of other aspects. In that respect CAAD reproduces manual drawing practices like overlay drafting.

Such practices in manual drawing are due to the limitations of the media used. The computer, however, cannot be treated as if it were merely a storage medium incapable of making explicit the implicit information that is rather effortlessly recognized in conventional architectural drawings.

As we can see in all branches and applications of computer science one of the primary merits of the computer is the ability to recognize and make explicit the information that is implicitly encoded in a description. For example, a computer vision system can transform the pixels of a digital image into a collection of parts and recognize in this collection all kinds of objects, even when they are partially obscured [Brooks 1981]. Given such ability of the computer to reproduce the performance of human perception and cognition even if on a limited scale, I find it impossible to endorse the opinion that a CAAD system should require explicit input of all relevant aspects of a design. Such redundancy only decreases comprehension, coherence and comprehensiveness of the overall description and increases requirements in storage space and processing speed.

The dissertation proposes that the computer should be able to recognize automatically descriptions of one aspect of a design in descriptions of other aspects, on the basis of domain knowledge that is integrated in architectural representations. This means that storage is reduced to the absolutely essential aspects, those that condense the most information and those that cannot be inferred from others. The main reason for automated recognition, besides economy of storage and reduction of input data, is that if the computer is to be a true aid to architectural design it should be able to recognize the same entities as its user and arrive at the same levels of abstraction with as little assistance and guidance by the user as possible. Only then we may expect a significant contribution of the computer to the quality and efficiency of architectural design.

In chap recognition of spatial primitives is an automated process which accepts as input digitized images of conventional architectural plans. In that respect the process outlined in this chapter bears similarities to the automated conversion of conventional drawings into computer documents, as attempted mainly in the fields of electrical and mechanical engineering [Hofer-Alfeis 1987; Joseph 1989]. Here too input is in the form of digitized images which are first vectorized, i.e., transformed into collections of picture element chains. These vectors are then categorized according to their type (straight line segments, arcs, circles, alphanumeric strings) and transformed into the appropriate CAD data structures. At this point similarities with our problem end. Once the digitized image is transformed into a collection of instances of CAD primitives, it follows that conversion is successful and complete. This means that the

ultimate effectiveness of the conversion can be measured by the descriptive effectiveness of CAD primitives. If the primitives are inadequate for representing the depicted objects by themselves, as is the case with most current CAAD systems, then the conversion, however successful from a technical perspective, is of little value without additional processing of the CAD primitives into primitives which represent the more abstract real-world entities that also concern design.

Automated conversion of engineering drawings into computer documents has been more successful in schematic drawings. These consist mainly of component symbols, their alphanumeric labels and connections between components (wiring). The limited number of these basic categories and the overall simplicity of the atomistic structure of schematic diagrams have proven of great assistance in converting conventional schematic drawings into CAD documents [Groen & van Munster 1984, 1986]. In mechanical engineering drawings conversion has been generally less successful because depicted objects may belong to a wider range of categories and may appear in unconventional or even arbitrary projections. Recognition of symbols is normally performed through techniques such as tracing closed loops of line segments [Joseph 1989].

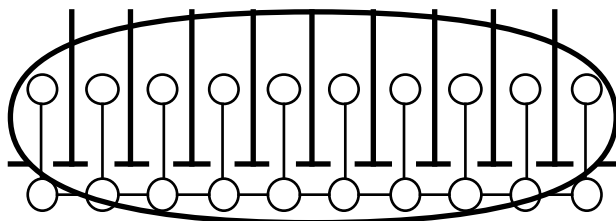
Recognition of spatial descriptive primitives in architectural plans cannot rely on the comprehensiveness of CAAD primitives. These rarely if ever include spatial primitives which, when needed, are normally approximated by the set of bounding construction elements. Therefore, the process presented in this chapter has to go beyond conversion of conventional architectural into CAAD documents. In fact, the approach is in principle applicable not only to digitized images of conventional architectural plans but also to CAAD documents.

The fundamental difference between the approach that underlies recognition of spatial descriptive primitives in chap and conventional CAAD systems is that the user is allowed to input drawings flexibly, practically as if they were analogue images, that is, in very much the same way he might draw them on paper. Organization of the drawn information into data structures such as layers, objects or frames, or into descriptions of design entities such as interior spaces, is left to the computer. Thus the computer can become a truly intelligent design assistant instead of just another medium for recording information. Automated recognition of spatial descriptive primitives is a fundamental first step towards the development of such intelligent CAAD environments which do not demand explicit input of all aspects of a design.

## 2.1.2 Locations

The spatial primitives of *chap*, that is, the atomic parts of descriptions of architectural plans in *chap* are *locations*. The term is derived from theoretical geography and economics [Alonso 1964, 1975; Haggett 1966; Isard 1956] where locational analysis is used to study the distribution of data such as actors, activities or equipment. A location in a building could be accordingly defined as the atomic parcel of space where a distinct activity can take place. The factors that determine the definition of a location in practice include the local and global nature and geometry of space boundaries, the allocation of actors and activities in a building, functional differentiations and elements of space perception.

I shall not attempt to elaborate a definition of locations in architecture as this is inhibited by the scaling-down assumptions of *chap*. A complete definition should take into consideration both the spatial articulation of form and the functional (operational) content of a design. In *chap*, however, only metric characteristics of architectural plans are considered. As a result, locations in *chap* coincide with physically defined parcels of space, i.e., the subdivisions of space determined by the building elements of the plan. This conscious oversimplification of the definition of a location in *chap* does not reduce the validity of the underlying approach nor of the resulting descriptions of architectural plans because *chap* is open to more elaborate definitions of locations. Any such definition simply implies that the locations of *chap* may have to be subdivided into more components (cfr. Figure 2-1).



**Figure 2-1** *An illustration of the distinction between interior spaces and locations  
each node of the access diagram represents a location*

The term ‘location’ is preferred to the more familiar ‘interior space’ which is used extensively in CAAD and the more vague ‘room’, although both would be in principle suitable for *chap*. The term ‘locations’ is adopted (although not properly used) in the dissertation as a reminder of the fact that, although spatial primitives in *chap* are defined by formal criteria only, they are in fact defined by the

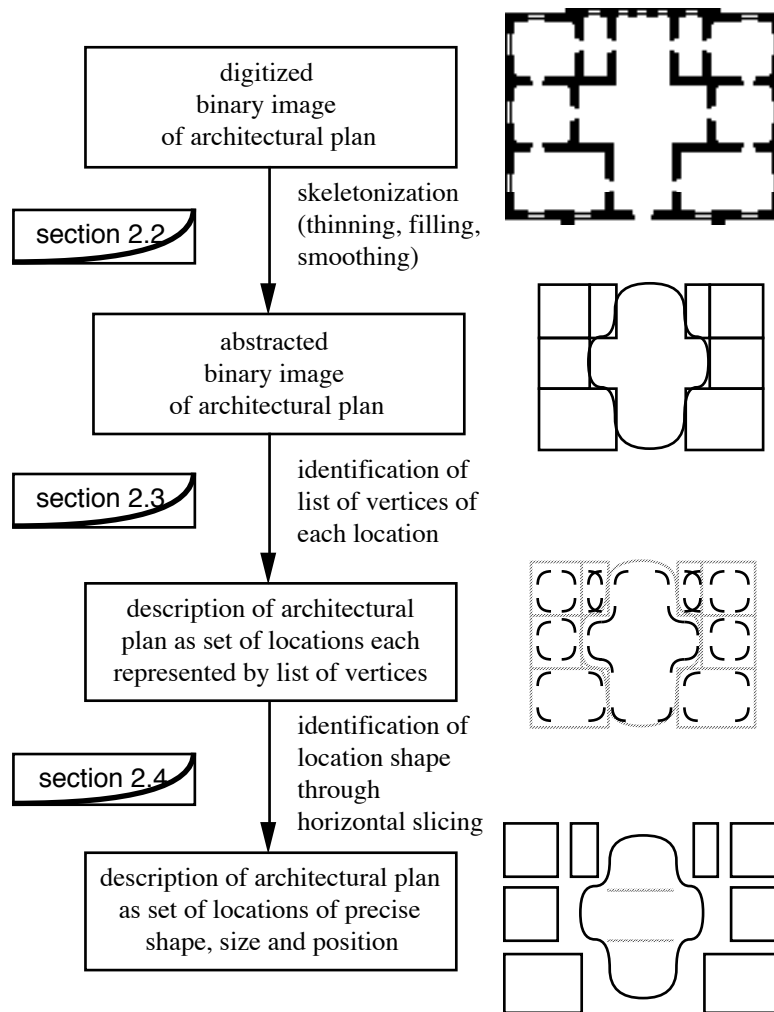
combination of formal and functional aspects. An interior space may contain more than one locations, as in the case where the dining and living rooms are unified into one larger room. Subdivision of an interior space into more than one locations also occurs in cases where functional differentiations are not so obvious: in Figure 2-1 each room is considered to comprise one location only, except for the corridor which is subdivided into a number of locations in accordance with the positioning of the entrances of other rooms. Such subdivision is not encountered in CAAD approaches and techniques that concentrate on formal aspects but has been employed in studies of pedestrian circulation in buildings [Tabor 1976b].

## 2.1.2 An outline of the chapter

Section 2.1 discusses the reasons for the use of spatial descriptive primitives in *chap* and for the automated recognition of these primitives. It also introduces locations, the primitives used in *chap* and explains the consequences of the scaling-down assumptions of *chap* for the definition of locations.

Section 2.2 considers the only part of *chap* that is not explored in detail, input of architectural plans in machine environment. This includes two basic substages, digitization and preprocessing. Both are essentially implementation-dependent, that is, they relate to the particular computer environment where *chap* may be implemented and therefore belong more to the development stages that should follow the dissertation. Section 2.2 specifies input of architectural plans mainly with respect to two points only: (a) digitization should be through optical digitizing devices and (b) the output of preprocessing should be a skeletonized image that corresponds to the essential description of the spatial articulation of an architectural plan, to the description that amounts to the rough sketch of a plan one produces from memory.

Section 2.3 considers precedent approaches to the recognition of locations in architectural plans. One approach that has been implemented in CAAD is through line following. Another approach which has been applied to images similar to abstracted architectural plans is based on the identification and correlation of the vertices of a shape.



**Figure 2-2** A schematic outline of the module of *chap* described in Chapter 2 with respect to the general outline of Figure 1-1

An implementation of the vertex-based approach is presented in sections 2.4 and 2.5. Section 2.4 describes a typology of location vertices that may be encountered in *chap* and the procedure used for the identification of the vertices of each location in an architectural plan. Recognition of the shape of a location on the basis of the list of its vertices is discussed in section 2.5. Given the quite large number of shapes possible even in the scaled-down environment of *chap* and their variations due to scaling, rotation and reflection, recognition of shape is based on decomposition into simple rectangular components. The relationships between these components amount to a definition of shape, as in generalized cones and similar descriptive techniques. Shape recognition through decomposition is helpful for the abstraction and classification of shape types because, by putting emphasis on the relationships rather than on the components and related features of a shape, it is possible to ignore local perturbations and geometric transformations of the same basic shape type.

The chapter concludes with section 2.6 which summarizes conclusions and outlines the extensions of the

techniques suggested in sections 2.4, 2.5 and 2.6 with respect to the ability of the proposed process for the recognition of locations to integrate various shape typologies and to extended beyond the scaled-down environment of CHQP, as well as its connections with some human perceptual mechanisms.

Figure 2-2 is a schematic outline of the process described in this chapter.

## 2.2 Input of architectural plans in machine environment

### 2.2.1 Digitization

A common technique of inputting existing drawings in machine environment is through manual digitizers, such as graphic tablets. In this, the user describes the drawing in terms of instances of the primitives of the system he uses by selectively tracing the lines of the drawing with a pointing device. The technique amounts to an interactive combination of vectorization and transformation of the vectors into the appropriate geometric primitives (cfr. subsection 2.1.1).

The absence of systems for the automated conversion of architectural drawings into CAAD documents means that there is no commercially available alternative to this technique. And as far as it concerns large, detailed construction or presentation drawings, the expectations for automated conversion of architectural drawings cannot be very optimistic. On the other hand, CHQP addresses levels of higher abstraction where detail and precision are not so binding. For these levels the currently available technology of optical digitizers (scanners) and related image processing programs suggests an approach to digitization that does not require the heavy workload and burden of interactive interpretation of a drawing through a manual digitizer.

Input of architectural plans into machine environment through optical digitizers results into the transformation of a conventional black and white architectural plan into a *binary image*, i.e., a two-dimensional array of pixels. Each pixel of the array has the value of either 0 or 1 (0 corresponds to white and 1 to black), hence the term 'binary'. Grayscale and colour digitizers are capable of also capturing shades of gray and colours but this added luxury is not considered in the dissertation because (a) architectural plans are mostly black and white and (b) grays and colours are visualization aids that do not add to the information on the spatial articulation

of a building provided by a plan.

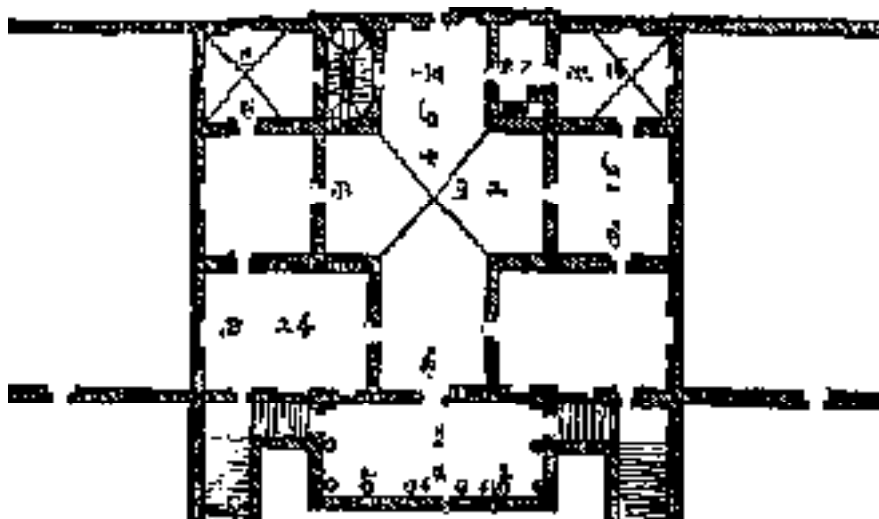
The binary image returned by an optical digitizer is in a more 'raw' state than the output of an interactive transformation of the plan into a CAAD document. Still, the binary image is sufficient if the only purpose of digitization is to store the drawing into a computer. Even if the drawing is used for some design task where partial modifications are required, optical digitization is still preferable because only parts of the binary image would have to be interactively transformed into complete CAAD documents.

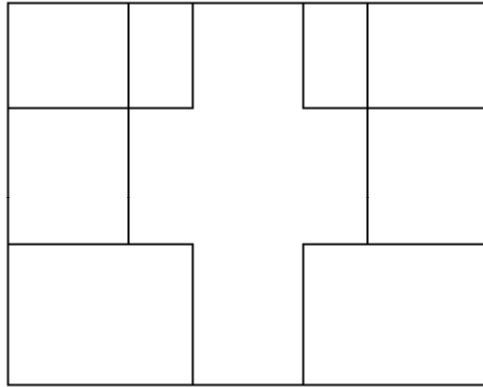
For the purposes of chap the low level of the information that is explicit in the binary image of an architectural plan offers a major advantage as it allows transformation of the image into descriptions that is not restricted by the limitations of the primitives of a CAAD system.

## 2.2.2 Preprocessing

Most optical digitizers can produce a quite accurate, detailed and low-noise binary image of an architectural plan. The handheld device used for the production of Figure 2-3 with its maximal resolution of 300 dots per inch represents the low end of optical digitizers but still manages to produce a readable copy with just a few local deformations.

Orthogonality constraints and the accuracy of current optical digitizers reduce the need for noise reduction, corrections, adjustments or other enhancement of the images of chap. Still, preprocessing of the binary image is required for its simplification: removal of information that is redundant or unnecessary for the recognition of locations facilitates operation of the procedures of chap.





**Figure 2-3** Andrea Palladio, *Villa Foscari, Gambarare Di Mira* (1669–60): digitized image of the architectural plan (top) and abstracted binary image (bottom)

The first part of preprocessing is a *smoothing* process which simplifies the scanned image into an *abstracted binary image* (Figure 2-3), a smooth (without local discontinuities) skeleton of unity thickness, by performing the functions of *filling* and *thinning*.

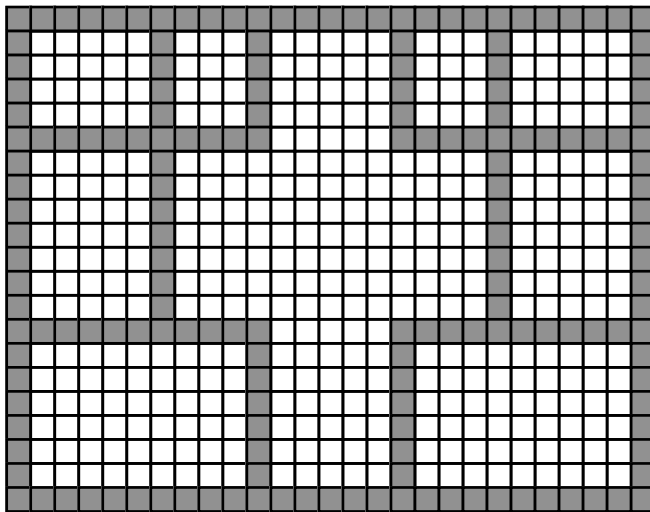
The smoothing process is similar in structure and purpose to skeletonization in optical character recognition whereby information which is not essential to the recognition of the character but simply result from the particular typeface or font, such as line weight and serifs, is removed leaving behind a skeleton which can be analyzed into a set of strokes and junctions which is characteristic of the character regardless of font or typeface. In chap the smoothing process produces an abstracted binary image (Figure 2-3) that contains only part of the total amount of information represented in an architectural plan. Vertical partitioning elements (including openings) are reduced to black lines of unity thickness, while material identifications, dimensions, verbal labels, stairs, and other secondary information, including vertical partitioning elements that do not belong to the boundary of a location, are removed. What remains is information about the boundaries and hence the position, size and shape of the locations. This information is necessary and sufficient for the recognition of each location and further on for the recognition of the structure of the whole plan.

The abstracted binary image bears strong similarities to the rough sketch one would produce from the memory of an architectural plan. These similarities suggest that the abstracted binary image can be considered as the essential representation of the spatial articulation of an architectural plan, the outline on which more elaborate descriptions are based.

Classical architectural plans exhibit certain characteristics that facilitate smoothing. Openings in a classical building are relatively small, there are few walls that do not bind a location and the boundary of each location

is normally quite well defined. Even if the boundary is formed in part or whole by a colonnade, the relation between the overall colonnade length and the distance between adjacent columns is such that allows the transformation of the colonnade into a line of unity thickness. The orthogonality constraints of CHQP also facilitate smoothing as they allow only two directions for the lines contained in the abstracted binary image.

The smoothing process has not been fully implemented in CHQP because it is largely dependent on the implementation environment. A preliminary investigation of skeletonization, smoothing, filling and thinning procedures [Rosenfeld & Kak 1982, pp.232–240; Pavlidis 1982, pp.199–208] suggests that production of abstracted binary images as perfect as the one in Figure 2-3 is to be expected not only within the totally orthogonal environment of CHQP but also when orthogonality constraints are relaxed.



```

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1
1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 1 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

**Figure 2-4** The abstracted binary image of Villa Foscari in array

form

*lack of precise metric correspondence between the array and Figure 2-2 is due to the coarse sampling grid used to reduce the number of picture elements for presentation purposes*

The second part of preprocessing is the removal of the white borders which inevitably (for technical reasons) surround the abstracted binary image. This task is performed by a procedure which removes all empty rows of the bit array of the abstracted binary image, starting from the first and last one and ending either way with the first encountered non-empty row. The same is then applied to columns. The array returned by this procedure (Figure 2-4) is reduced in size, thus reducing storage and search space. For example, the perimeter of the plan (partitions which separate the exterior from interior spaces) lies mostly in the first and last columns and rows of the array and this, as we shall see in the section 2.4 is used extensively in the recognition of the vertices of each location.

## 2.3 Approaches to the recognition of locations in architectural plans

### 2.3.1 Line following

The linear appearance of architectural plans suggests that *line* or *contour following* techniques are applicable to the recognition of locations bound by the lines. Line following techniques were employed in the few cases where automated recognition of spatial primitives in architectural plans has been considered [Hall 1983; Lawson & Riley 1982]. A simple line following algorithm for the abstracted binary images of Figures 2-3 and 2-4 would be:

- 1 Select a random edge junction in the plan (Figure 2-5a).
- 2 Select the first edge counterclockwise (Figure 2-5b).
- 3 Move along this edge until a new edge junction is encountered.
- 4 Select the first edge counterclockwise (Figure 2-5c).
- 5 Repeat steps 3 and 4 until the initial edge junction (of step 1) is encountered. The loop of edges followed is the boundary of a location (Figure 2-5d).
- 6 Select the first edge junction where a choice between different edges was made in the recognition of the previous location (Figure 2-5e).
- 7 Repeat steps 2 to 5 to find the boundary of the next

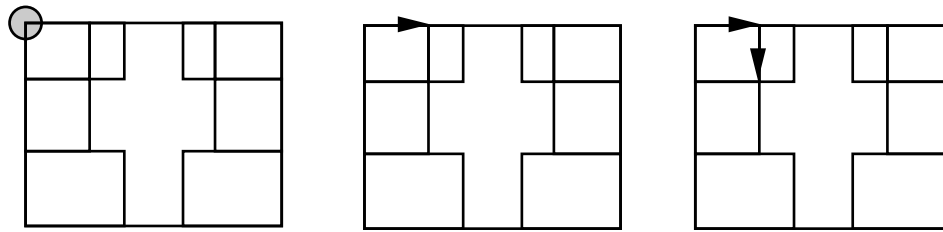
location. The only difference in the rules concerns the selection between different edges at a junction: if this junction has been already encountered, edges which have been previously selected *in the same sense of direction* are ignored (Figure 2-5f).

- 8 Repeat steps 6 and 7 until all edges have been selected and therefore all locations have been recognized (Figure 2-5g). In selecting a starting point for a new location (step 6), if an edge junction does not offer any previously unselected edges to choose between, then backtrack to the location recognized before. (An instance of this case occurs after the recognition of the seventh location, when it is necessary to backtrack as far as the third location to find an edge junction suitable as the starting point for the eighth location.) If this does not yield any results and there still remain unselected edges in the plan, then look for unencountered edge junctions on the unmarked edges (this implies a case of inclusion).

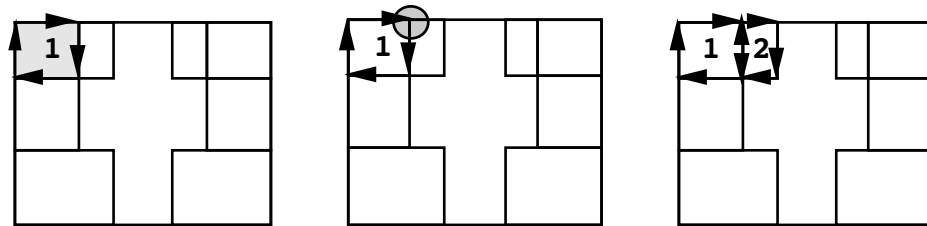
This simple line following technique can successfully identify closed loops of partitioning elements in an abstracted binary image and thus the position and even the shape of the location bound by the loop. However, the effectiveness of the technique should not obscure a number of disadvantages that restrict the applicability of line following in *Chop*. These include general algorithmic problems of line following and, perhaps more significantly for the dissertation, the lack of distinction between the *boundary* of a location and the location itself.

The disadvantages of line following techniques in general include the high degree of algorithmic and computational tenacity required for the recognition of even simple shapes [Duda & Hart 1973, pp.290–293; Joseph 1989; Nevatia 1982, p.139]. Signs of this tenacity can be found even in Figure 2-4. Each selected edge must also be marked with the sense of direction the line follower moves along, so as not to disqualify the inclusion of the edge in the boundary of another location. It is also necessary to keep a record of every edge junction where choice between different edges was made for subsequent investigations and update the record every time the line follower passes again from one of the vertices included in it.

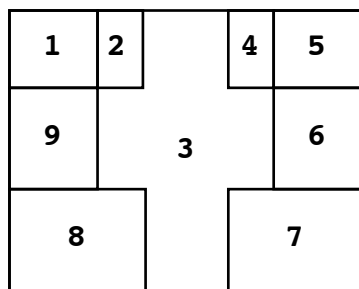
Line following is intrinsically a serial process which transforms shape recognition into something similar to solving maze puzzles, completely unlike human perception. Aside from that an error made at any step most probably means that succeeding steps will also be in error [Duda & Hart 1973, p.292], line following does not facilitate the integration, recognition and exploitation of more general constraints which lead directly to higher levels of abstraction.



a (left): an edge junction is selected at random;  
 b (middle): the first edge counterclockwise is selected;  
 c (right): at a junction the first edge counterclockwise is selected;



d (left): when the starting edge junction is encountered, a location has been identified;  
 e (middle): the next starting edge junction is the first junction where choice between different edges was made in the recognition of the previous location;  
 f (right): the same steps are used to recognize the next location; the only difference lies in the selection of edges at junctions which have been previously encountered

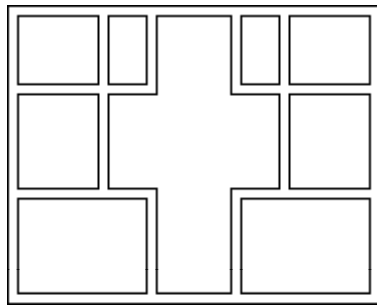


g: results of the procedure; numbers denote the order of recognition of each location

**Figure 2-4** Recognition of locations through line following

Another disadvantage of line following that is particularly significant for the dissertation and chap is the lack of distinction between the boundary of the location and the location itself, i.e., between the ‘solid’ and the ‘void’ parts of an architectural plan. The duality of solid building elements and voids which accommodate the activities in the building is considered to be an issue particular to architectural design [Steadman 1987, p.246]. Steadman suggests that, since the two aspects are complementary and an architect moves continuously between the two in designing, the distinction is useful for CAAD presumably because it allows more comprehensive consideration and treatment of architectural form. As suggested by the sum-of-parts and the multiple / holistic representation fallacies (subsections 1.4.2 and 1.4.3), this distinction is not a general case in CAAD. In the few cases where it is made the advantages which emerge are quite obvious, as in shape

grammars [Gips 1975; Stiny 1975, 1980; Stiny & Mitchell 1978a, b] and rectangular arrangements [Mitchell et al 1976; Steadman 1983].



**Figure 2-5** *Void modelling: the floor plan of Villa Foscari*

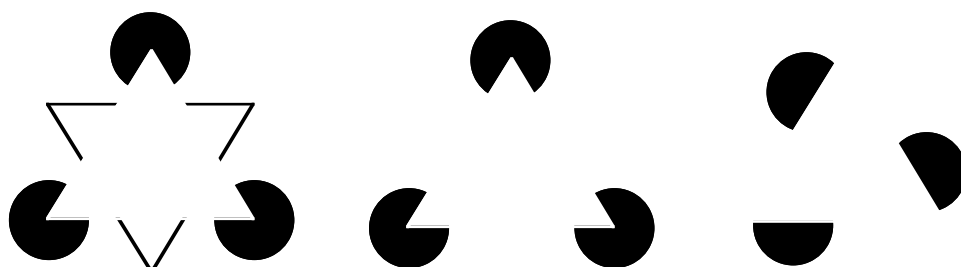
An example of successful distinction between ‘solids’ and ‘voids’ that is relevant to the recognition of locations through line following is *void modelling* [Yessios 1987]. Yessios substitutes the primitive solid volumes of solid modelling with void ones which are obviously better suited to the description of buildings as containers of space. The use of void volumetric primitives in an architectural plan means that vertical partitioning elements are symbolized by double lines which bound on the one hand the interior spaces and on the other the overall area of the plan. The space between the two boundaries is occupied by the solid components of the building (Figure 2-5).

This form of distinction between locations and the building elements that bind them allows the description of both ‘solids’ and ‘voids’ at the same time and subsequently leads to fewer inconsistencies with the human interpretation of architectural plans and also to fewer operational problems. For example, in Figure 2-4 locations 7 and 8 are described as if having five edges even though they have rectangular shape. To correct the artificial subdivision of the one edge further processing is needed. In Figure 2-5 both locations can be immediately identified as having four edges, as each location is bound by a distinct polygon.

Techniques such as void modelling resolve several conceptual and operational problems that arise from the lack of distinction between a location and its boundary but cannot negate the general disadvantages of line following. The contribution of void modelling to the dissertation is in the manner it distinguishes between the boundaries of building elements and the boundaries of locations. A similar distinction is followed in chap (cfr. section 2.4). This distinction is, however, more related to an alternative to line following, recognition of a shape on the basis of its vertices and their connections.

### 2.3.2 Vertices in human and computer vision

Line following techniques are generally justified by the undeniable importance of edges to the human visual system. However, by sticking to edges only we underestimate the significance of critical areas in an image, such as edge intersections and their correlation to constraints imposed by the nature and the meaning of the image. A perfect example of this point are *illusory* (or subjective or cognitive) *contours* [Kanizsa 1955, 1974] where the shape, size and position of a figure in an image are perceived without perception of its actual contour.



**Figure 2-6** *Illusory contour figures, I*  
left (a): a strong case; centre (b): a weaker case; right (c): an even weaker case

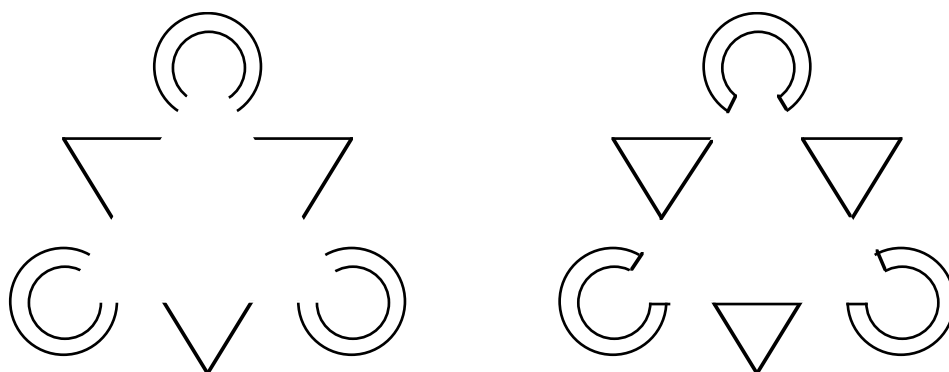
In Figure 2-6a the stimulus, three incomplete black disks and three pairs of edges connected at an angle, supports perception of the image as a triangle that occludes another triangle and three complete black disks. The occluding contour of the occluding triangle is in fact invisible; its presence is deduced by correlating parts of the stimulus which are reminiscent of occlusion conditions. Still, the preference for the perception of the illusory contour triangle is so strong that the area supposedly occupied by it appears brighter than the rest of the continuous white area in the image. This difference in brightness would only be natural in the case of a white triangle in front an equally white background.

Illusory contour perception is supported for both regular / familiar and irregular / unfamiliar figures. However, not all illusory contours are perceived as readily as Figure 2-6a. Figure 2-6b is weaker, and Figure 2-6c even weaker. The difference between 2-6b and 2-6c suggests that the mere possibility of an occluding contour is not sufficient for its perception. Also, perception of illusory contours is not dependent upon the presence of fragments of the contour in the stimulus. Figure 2-7a supports perception of an illusory triangle, while 2-7b does not, although one could say that more parts of the edges of the illusory triangle are

present in Figure 2-7b than in 2-7a. These ‘parts’, however, cannot be structured together to form the boundary of an illusory triangle but form edges of other shapes. By contrast, in Figure 2-7a the elements which so strongly support the presence of an illusory triangle are not edges themselves but belong to the category Marr [1982, pp.69–79] calls *terminations*, that is, discontinuities in edge orientation or ends of edges. Terminations are characteristic of occlusion conditions and the alignment of terminations along a hypothesized boundary, that is, the homogeneity and the type of spatial arrangement of the stimulus, is what supports perception of an illusory contour.

A final point with particular relevance to our problem is that perception of the vertices of a shape is a necessary and sufficient condition for the perception of the whole illusory contour figure (Figure 2-6b), while perception of its edges only does not necessarily support perception of the whole (Figure 2-6c).

The significance of vertices in human vision is transferred to the computer by the approach to the recognition of line drawings of objects with trihedral vertices initiated by Guzmán [1968], elaborated by Huffman [1971] and Clowes [1971] and algorithmically formalized by Waltz [1975]. This approach relies on the identification of the type of edge junctions, that is, of trihedral vertices of objects in the scenes and of points where boundary edges occlude other edges.

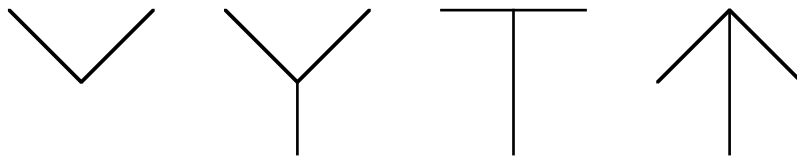


**Figure 2-7** *Illusory contour figures, II*

*left (a): an illusory contour figure, right (b): a figure that does not support illusory contour perception (after Kanizsa [1974])*

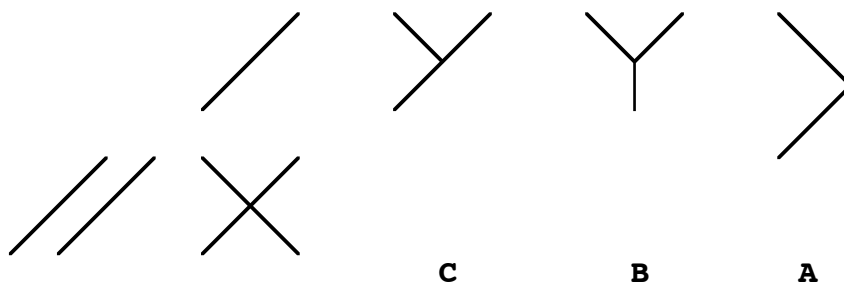
One of the main advantages of the approach is that recognition of the scene structure is based a limited number of junction types. If we exclude cracks and shadow edges from scenes of objects with trihedral vertices, there are just eighteen junction types [Winston 1984, pp.51–55]. These types are furthermore produced by the labelling of the edges of the four essential junction categories (Figure 2-8) with respect to whether the edges are concave, convex or part of

the boundary of an object. By identifying the type of each edge junction and the junctions it is physically connected to, it is possible to identify not only the faces of the objects in a scene but also their orientation and other attributes.



**Figure 2-8** *The four basic categories of edge junctions in line drawings of objects with trihedral vertices*

A variation of this approach is proposed by Walters [1986]. On the grounds that certain edge junctions receive preferential enhancement in human vision, Walters suggests a similar selective enhancement hierarchy for computer vision. This hierarchy (Figure 2-9) is dissimilar from the types of Figure 2-10 in that it discriminates explicitly between 'end' and 'middle' edges, that is, between the occluding contours and the remaining edges of an object. Despite this, in both cases the information required for the recognition of the structure of a scene is localized on edge junctions. A typology of these junctions and their connectivity are the essential tools for building a basic description of the scene.



**Figure 2-9** *A hierarchy of preferential stimulus enhancement order of enhancement (descending): A—B—C (after Walters [1986])*

Recognition systems based on the identification of the vertices of a shape and their connectivity are more attractive than line following for  $\text{Chap}$ . Vertex based recognition is essentially parallel and hence allows greater tolerances and less tenacity in its implementation. This means that abstracted binary images less perfect than Figure 2-3 can be used for  $\text{Chap}$  without extensive noise reduction, enhancement or other preprocessing. Also, by concentrating on the vertices of a shape, it is possible to distinguish between the location and its boundary not only on a conceptual but also on the operational level (cfr. section 2.4). Finally, while in line following general shape constraints have to be integrated into the basic algorithm of subsection

2.3.1, recognition of a shape on the basis of its vertices relies on an explicit structure of such constraints. Such explicitness is instrumental for the evaluation and augmentation of rules derived from domain constraints and hence for the development of an effective, efficient and comprehensive recognition system which can moreover serve as a platform for the analysis of architectural perception.

## 2.4 Identification of the vertices of a location

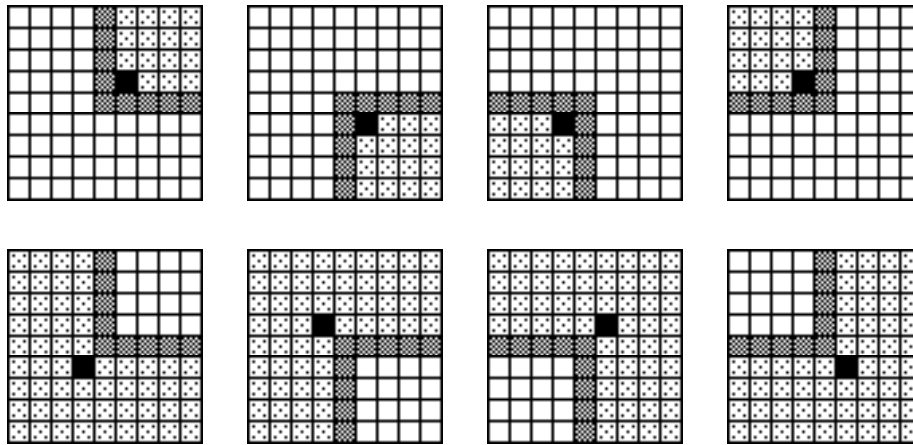
The significance of vertices in human vision and the successful use of vertex-related constraints in computer vision led to the development of a recognition technique based on the identification of the vertices of each location, as the scaling-down constraints of *chap* allow the definition of vertex and connectivity typologies as compact as the ones used by Waltz [1975]. These typologies form an effective set of rules that suffices for the recognition of the essential aspects of each location in abstracted binary images.

By 'vertices of a location' I refer not to the vertices of the boundary but to the vertices of the bounded area. In Figure 2-4 the top left vertex of the top left location is not on the first row, first column but on the second row, second column. This is an effective way of discriminating between the 'solids' and the 'voids' in an architectural plan without being implicated in the computational and perceptual problems associated with techniques such as void modelling, simply by exploiting the fact that edges in an abstracted binary image (and in an architectural plan in general) are *line edges* according to the classification by Herskovits [1970], i.e., they represent abrupt changes in brightness in both directions. Subsequently edges in binary images can be treated as spatially well-defined regions, completely distinct from the white regions they bound.

### 2.4.1 A typology of vertices

The orthogonality constraints of *chap* can be systematized into a reliable typology of location vertices which consists of the eight possible types of orthogonal location vertices: I, II, III, IV,  $\neg$ I,  $\neg$ II,  $\neg$ III and  $\neg$ IV (Figure 2-10). The names denote which quadrants of the Cartesian coordinate system, the origin of which is the location vertex, are occupied by the location the vertex

belongs to. For example, type III means that the location the vertex (origin) belongs to occupies the third quadrant and type  $\neg$ II that the location occupies all but the second quadrant.



**Figure 2-10** *The eight vertex types in the orthogonal environment of chap*  
top row (from left to right): types I, II, III, IV; bottom row:  $\neg$ I,  $\neg$ II,  $\neg$ III,  $\neg$ IV;  
dark cells represent black pixels (location boundaries); light cells represent white pixels which belong to the location under investigation; black cells represent the vertex under investigation

Knowledge of the type of a vertex of a location contains implicit knowledge of the directions in which the two vertices connected to it lie. In the orthogonal environment of  $\text{chap}$  each vertex is defined by two edges vertical to each other. At the end of either edge there is another vertex, which is defined by an edge of the previous vertex and another edge. The new edge can only be vertical to the other (at either  $90^\circ$  or  $270^\circ$ ) and therefore, the new vertex may be only of either of two types. It follows that for any given vertex type we can form positive expectations for the type of the two vertices connected to it: each may belong just to either of two types (Figure 2-11).

The type of a vertex can be determined through the use of the eight array templates (windows) in Figure 2-10, each of which identifies uniquely a vertex type. The size of these windows is related to the tolerances for local perturbations, gaps and similar imperfections of the abstracted binary image. For example, transposition of an edge by a single pixel does not result into a vertex of a location in an architectural plan.

The specificity of the relationships between vertices of a location implies that identification of a location (or rather of its vertices) essentially amounts to identification to one of its vertices. Although we could start from a vertex selected at random, it is preferable to use a standard and safe technique to establish a starting point. Such a technique can be based on the relations described in Figure 2-11b which concern

special cases of vertices. For example, the top left vertex of any location, regardless of row- or column-major order in search, is invariably (due to orthogonality constraints) of type II. This vertex belongs to the second row or column of the array, since in the final stage of preprocessing all empty rows and columns were removed.

vertex	up	right	down	left
<b>I</b>	II,¬III	IV,¬III	*	*
<b>II</b>	*	III,¬IV	I,¬IV	*
<b>III</b>	*	*	IV,¬I	II,¬I
<b>IV</b>	III,¬II	*	*	I,¬II
<b>¬I</b>	III,¬II	III,¬IV	*	*
<b>¬II</b>	*	IV,¬III	IV,¬I	*
<b>¬III</b>	*	*	I,¬IV	I,¬II
<b>¬IV</b>	II,¬III	*	*	II,¬I

first	second	second-last	last	vertex
<i>second and second-last row:</i>				
II	III	II	III	
<i>second column:</i>				
II	I	II	I	
<i>second-last column:</i>				
III	IV	III	IV	

**Figure 2-11** Vertex connectivity

top (a): general relations: vertex types connected to each vertex type; asterisks denote that no connected vertex may be expected; bottom (b): special cases: the types of the first and last two vertices of the second and the second-last column and row

## 2.4.2 Identification of the vertices of a location

On the basis of the above vertex typology and connectivity, the following procedure is used for the identification of the vertices of each location:

- 1 The first vertex of type II encountered in the second row of the abstracted binary image array is marked as the starting point of the first location, i.e., is assigned a value specific to this location (unique identification — Figure 2-12a).
- 2 The next two vertices (the ones connected to the starting point) are found according to positive expectations formed on the basis of vertex connectivity (Figure 2-12b). Then the vertices connected directly and indirectly to each of these two are found and marked for identification purposes, as well as in order to limit the search space for the identification of the remaining locations:
  - 2.2 If the limits of the image are encountered either horizontally or vertically, both vertices as well as the





```

1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1
116 0 0 016 113 0 0 013 0 0 013 0 0 013 117 0 0 017 1
1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1
118 0 0 0 0 0 0 018 1 0 0 0 0 0 119 0 0 0 0 0 0 019 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1
118 0 0 0 0 0 0 018 113 0 0 013 119 0 0 0 0 0 0 019 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

*e:when no unmarked vertex of type II remains, then all vertices of all locations have been identified*

**Figure 2-12** Identification of the vertices of each location

```

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
111 0 0 011 112 012 113 0 0 013 114 014 115 0 0 015 1
1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 1
111 0 0 011 112 012 1 0 0 0 0 0 114 014 115 0 0 015 1
1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1
116 0 0 016 113 0 0 013 0 0 0 013 0 013 117 0 0 017 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1
1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1
116 0 0 016 113 0 0 013 0 0 013 0 0 013 117 0 0 017 1
1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1
118 0 0 0 0 0 0 018 1 0 0 0 0 0 1 2 0 0 0 0 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
118 0 0 0 0 0 0 018 113 0 0 013 1 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0

```

**Figure 2-13** Identification of vertices lying outside the perimeter of the plan

*Even if the bottom right location of Villa Foscari is removed, the top left vertex remains and is still a type II vertex. Subsequently, it is selected as a starting point for the identification of the vertices of another location which, in this case does not exist. The vertex is identified as lying outside the perimeter of the plan because it does not connect to any vertices to the right and down; instead, the limits of the array are encountered. Any vertex which does not connect to another vertex in either of the two directions we expect it to, as well as all vertices connected to it, are identified as vertices lying outside the perimeter of the plan and are marked with the universal value 2.*

The resulting sets of location vertices can be checked for accuracy by exploiting secondary properties of the vertex typology. For example, by assigning the value of vertices of type I, II, III and IV to 1 and the value of vertices of type ¬I, ¬II, ¬III and ¬IV to -1, the sum of the values of all vertices of a location is always equal to four.

The procedure of Figure 2-12 differs from Waltz [1975] in that

- the number of vertex types is much smaller in ChQP, as is the number of possible connections between

- vertices of different types
- the type of a vertex can be recognized directly from local information, without reference to the type of antecedent adjacent vertices
  - global shape attributes cannot be directly derived from the type of vertices

The first two differences imply that constraint propagation and relaxation on the scale used by Waltz are not required to determine the type of vertices in an abstracted binary image, while the third implies that constraint propagation among vertices is not sufficient for the recognition of shape types and makes Waltz's procedure inapplicable to our case. Apparently this is due to the level of definition of the vertex typology: the types used by Waltz represent higher-level primitives (i.e., closer to final goal of recognition) than the types used in chap.

At first sight the procedure might appear essentially identical to what could be achieved by a combination of line following technique with void modelling. Although this is partially true with respect to performance in a totally orthogonal environment, the difference is not merely an issue of talking about vertices instead of edges and their junctions. Techniques like void modelling may boost the performance of line following by removing ambiguities related to partitions (edges) and their junctions but the whole procedure remains essentially serial. By concentrating on vertices recognition of locations becomes essentially parallel, in very much the same way as Waltz's procedure: recognition is not achieved by the identification of closed loops but rather through the investigation of relations between critical features of an image. The potential of this approach becomes evident when we consider more complex floor plans where locations are not completely bound by closed wall loops. The parallel nature of the procedure has other advantages too, including comparatively high tolerances of error both in the input data and the execution of the algorithm. One of the consequences of this is that the procedure can be easily extended to cover architectural plans with oblique and curved lines (cfr. subsection 2.6.1).

### 2.4.3 Why the list of vertices of a location is not used directly for the identification of the shape of the location

The identification of all vertices belonging to a location is obviously not enough. The *shape type* of the location is also required, as an abstraction of the general spatial properties of the location and as a means of structuring the description of a

location. The technique which generally provides immediate results with respect to this problem is classification into a number of predetermined classes, each represented by a stored model of a pattern. New patterns are matched to the known ones (templates) and classified with the best match. In our case, the shape of a location can be determined by comparing it to a finite number of templates of prototypical shapes. By scaling and translating the templates to the scale and position of each location a simple overlay of the two patterns can be used to classify the location with respect to shape type (Figure 2-14). One major disadvantage of template matching is that, with the current computer and scanning technology, it is computationally expensive if patterns occur at different scales and orientations [Nevatia 1982, p.16], as in our case. Task-specific devices make template matching more efficient but even these are still rather limited with respect to translation and especially to scaling templates.

area match: 100% 67% 67% 67%  
 boundary: 100% 63% 50% 63%

Figure 2-14 Template matching for the recognition of the orientation of an L-shape (templates are represented by hatched shapes)

(II,I,IV, (II,I,IV, (II,I,¬III, (II,¬IV,II,  
 ¬II,IV,III) III,¬I,III) I,IV,III) I,IV,III)  
  
 (II,I,¬III,I, (II,I,IV,¬II, (II,¬IV,II,I,  
 (II,¬IV,II,I,  
 IV,¬II,IV,III) IV,III,¬I,III) ¬III,I,IV,III)  
 IV,III,¬I,III)

Figure 2-15 The four orientations of shapes in chap. the vertex type configuration is given in counterclockwise order, starting from the top left (in row-major order) vertex


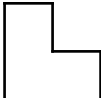
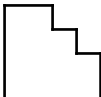
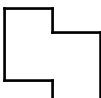
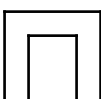
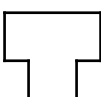
An alternative is to use not the patterns themselves but features extracted from the patterns, that is, to identify shape through classification of features, because feature selection and extraction are efficient manners of reducing the dimensionality of a pattern [Kittler 1986]. In our case the most easily recognizable features of an orthogonal shape that could guarantee a unique and efficient identification of shape type are the number and type of its vertices. These features

define a solution space within which shapes could be classified. However, subdivision of the solution space into clusters of similar patterns would be neither effective nor reliable because of the quite disparate alternative vertex type configurations a shape may assume due to the four possible orientations of each shape in CHQP (Figure 2-15). A way of combining the directness of templates and the abstraction of features is to consider the feature vector of each shape as a template in a predefined typology of location shapes. This technique generally works better with a limited number of types, for example shapes of eight vertices or less. Shapes of more than eight vertices (Figure 2-16) are characterized by an extremely low frequency in architectural plans and therefore can be treated as variations or subtypes of just a few types.

Figure 2-17 depicts the possible configurations of vertex types for the six shape types of eight vertices or less. These configurations can be used as templates in a simple yet efficient rule system which checks the number of vertices, the vertex types included, the number of instances of each type and finally the precise order of each instance in the configuration, i.e., which vertex type is connected to which.

While this technique works well with the limited number of shape types in Figure 2-17, problems arise if more shapes, such as those in Figure 2-16, are added as new types or even as subtypes or variations of the initial types. Augmentation of the shapes that can be recognized by the procedure through the addition of new types may result into an explosion of the rule database size, regardless of the sophistication of rule control mechanisms. If the new shapes are considered as subtypes, for instance in the sense of Figure 2-16, a reorganization of the rule system is necessary as the number of vertices is of little help for such cases. The configuration of vertex types is more useful because we could obtain the configuration which characterizes one of essential shape types by subtracting a number of vertices from a more complex shape. This, however, means that the vertices are implicitly combined into higher level entities. An explicit transformation of the list of vertices into higher level components of the shape is therefore preferable and is in fact used for the recognition of location shape in CHQP (cfr. section 2.5).


**Figure 2-16** Shape types of ten and twelve vertices as subtypes of shape types of eight vertices: a first approximation  
 left column: shapes with eight vertices (the basic shape types); middle column: shapes with ten vertices; right column: shapes with twelve vertices; broken outlines indicate multiple instances, i.e., shapes that can be considered as subtypes of more than one basic types

	II, I, IV, III
	II, I, IV, III, $\neg$ I, III or II, $\neg$ IV, II, I, IV, III or II, I, $\neg$ III, I, IV, III or II, I, IV, $\neg$ II, IV, III
	II, I, IV, III, $\neg$ I, III, $\neg$ I, III or II, $\neg$ IV, II, $\neg$ IV, II, I, IV, III or II, I, $\neg$ III, I, $\neg$ III, I, IV, III or II, I, IV, $\neg$ II, IV, $\neg$ II, IV, III
	II, I, $\neg$ III, I, IV, III, $\neg$ I, III or II, $\neg$ IV, II, I, IV, $\neg$ II, IV, III
	II, I, IV, $\neg$ II, $\neg$ III, I, IV, III or II, I, IV, III, II, $\neg$ IV, $\neg$ I, III or II, I, $\neg$ III, $\neg$ IV, II, I, IV, III or II, I, IV, III, $\neg$ I, $\neg$ II, IV, III
	II, I, $\neg$ III, I, IV, $\neg$ II, IV, III or II, $\neg$ IV, II, I, IV, III, $\neg$ I, III or II, I, IV, $\neg$ II, IV, III, $\neg$ I, III or II, $\neg$ IV, II, I, $\neg$ III, I, IV, III

**Figure 2-17** Shapes of up to eight vertices and their vertex type configurations

*the first of alternative vertex type configurations is the one of the depicted instance while the others correspond to the possible rotations of the shape; vertex type configurations are given in counterclockwise order, starting from the top left (in row-major order) vertex*

The shortcomings of a shape recognition system based on Figure 2-17 exemplify the generally accepted problems of template matching and classification of features and in particular that these techniques are useful in cases where “the number of classes and the variability within a class are small” [Nevatia 1982, p.16], as in optical character recognition. In chap we are dealing with an infinite number of shapes which cannot be grouped into a relatively small number of classes on the basis of superficial features like the number and the type of its vertices.

Intuitive interpretation and categorization of such shapes relies more on general and/or abstract characteristics, such as direction and orientation of repeated elements, and reference to simpler shape types. There is experimental evidence that shape recognition is not based on precise boundary description. Recognition tests using relatively complex figures suggest that local variations and anomalies are not generally critical for the description of a shape, while “when that same physical region *constitutes* the form, there is no difficulty in perceiving and storing all its nuances” [Rock 1983, p.55]. It appears that whether a region constitutes part of the overall shape, or is inconsequential

with respect to, is a matter of part-whole relationships, relative size and focus of attention: “the perturbations characterize the mode of the contour but not the shape of the figure” [Rock 1983, pp.55–56].

In decision-theoretic pattern recognition, classification in cases of similar complexity relies more on second-order measures, such as the complexity of a shape, measured in terms of “wiggleness or jaggedness” (“total absolute curvature summed over the border”), compactness (perimeter squared divided by area), or amount of information required to specify a pattern (“presence of equal parts, periodicities, or symmetries” reduce complexity) [Rosenfeld & Kak 1982, p.265]. These and other measures of global shape properties, such as proportions, convexity or orientation are not considered to be appropriate for this stage of *Chap*. The reason is that they result into identification of membership to a general class and not into a complete and precise description of each location. This description is necessary for the investigation of spatial and geometric relationships between locations and the subsequent description of the whole architectural plan.

## 2.5 Recognition of shape through decomposition

Figure 2-16 illustrates that complex shapes can be produced from the simpler ones through the addition or subtraction of rectangular parts. Although the rules of the addition / subtraction seem to be either arbitrary (*ad hoc*) or very abstract and ambiguous, the whole process seems perfectly reasonable within the framework of the manipulation, composition and decomposition of parcels of space in architectural design. Such additive / subtractive procedures have been extensively used in CAAD: in space allocation and planning [Eastman 1975], shape grammars [Gips 1975; Stiny 1975, 1980; Stiny & Mitchell 1978a], rectangular arrangements [Mitchell et al 1976; Steadman 1983] and elsewhere, with emphasis on the purpose of the procedures and their results.

*Chap* concentrates on the addition / subtraction of instances of the most primitive location shape in a totally orthogonal environment, the rectangle, as a means of describing the structure of the shape of a location, that is, these general characteristics which prevail intuitive shape recognition and categorization. The description of each location as the product of addition / subtraction of instances of a basic primitive shape aims at representing the corresponding additive / subtractive procedures through

which locations are formed in architectural design. Representation of these procedures is instrumental in establishing the relationships between the final design product and design decisions which constrain and control these procedures. In *CHAP*, unlike shape grammars and similar techniques, this representation is in a declarative rather than procedural manner, that is, it does not aim directly at making explicit the sequence of design procedures through which a location was produced. Instead, *CHAP* attempts to provide a coherent general-purpose description of an architectural plan and its parts. This description offers a more comprehensive basis for the investigation of the process of architectural design than formalisms developed for the automated production of designs.

A basic reason for the description of locations through decomposition into primitive parts is that predefined typologies of location shapes are neither convenient nor helpful with respect to the investigation of architectural form in *CHAP*. Architectural design is characterized by combinatorial definitions of formal aspects on the basis of composition rules which cover relations between locations and parts of locations, as well as other constraints imposed on the shape of a location, for instance by construction or site topography. These constraints implicitly define typologies of possible configurations, similarly to structural pattern recognition, where pattern classes are also implicitly defined by rules of shape composition [Watanabe 1985, p.387]. But unlike structural pattern recognition, the classification of architectural form involves multiple, overlapping points of view and thus defeats direct classification, unless points of view are hierarchically embedded. On the other hand, the description of a pattern in unambiguous geometric and spatial terms forms a sound basis for interpretations and classifications from various points of view, including that of shape categorization.

### 2.5.1 Approaches to shape decomposition

A primary issue in the approach to shape description adopted in *CHAP* is the establishment of a reliable and comprehensive manner of decomposing a location into its constituent parts. This task is made no lighter by the current strong tendency in pattern recognition and computer vision towards descriptions of objects and scenes in terms of their parts [Brooks 1981; Minsky 1975; Nevatia 1982, p.55; Pentland 1986; Rosenfeld & Kak 1982, pp.55–56 and 276; Winston 1975]. Such *structural descriptions* attempt to capture the articulation and structure of an object or scene

and thus offer several advantages over simple classification schemes. These advantages include the ability to identify and compare partial differences and similarities between images, through which effective distinction between different classes can be made as small changes in the appearance of the object or scene cause only small changes in its perception by the machine [Nevatia 1982, pp.62–63; Winston 1975].

Structural descriptions include cases where a scene or an object is described in terms of its natural or canonical parts [Brooks 1981; Pentland 1986; Winston 1975] and cases where the parts are defined by resolution-dependent sampling, as in chain coding [Freeman 1961], or adaptive hierarchical subdivision, as in quadtrees [Samet & Rosenfeld 1980; Samet 1984a, b; Jenkins & Tanimoto 1983].

Natural parts assist the development of conceptually elegant and semantically rich representations but involve the fundamental problems of what constitutes a natural part, how it can be recognized and how it relates to other parts. It is only too easy to adopt commonsensical views which arbitrarily combine elements from different aspects. For example, one might criticize Winston's [1975] toy-block arches on the grounds of that spatial relations between the blocks are treated as if they were merely functional (support). Of course this criticism does not affect the overall quality of Winston's thesis, since the thesis mainly concerns issues other than providing the ultimate description of an arch, but reveals a rather simplistic view of the world that underlies many approaches to computer intelligence. Another example of this view in an otherwise elaborate and sophisticated context is Minsky's [1988] descriptions of chairs and arches (both monolithic and modular) where geometric, functional and construction aspects are indiscriminately and incoherently intertwined in a manner that undermines the effectiveness of their overall framework.

The main disadvantage of most attempts to describe objects or scenes in terms of their natural or canonical parts is that they do not go beyond superficial human knowledge in the sense that they do not analyse the very effective but at the same time quite abstract categorizations and decompositions we use in everyday life. These suffice for communication between humans but still offer incomplete or vague representations which are supplemented, corrected or made specific through complex underlying knowledge structures. These structures pass unnoticed in everyday life because they are general and basic instruments common to all who use the corresponding categorizations and decompositions. For machine intelligence the recovery of these knowledge structures is essential for making explicit the perspectives that are implicit in commonsensical views on the composition and articulation of scenes and objects and

hence for the meaningful manipulation of such views.

Other descriptive techniques are unambiguously artificial devices for lowering the dimensionality of a pattern. For example, chain coding is a way of describing a two-dimensional curve by a one-dimensional string (Figure 3-3). The modesty of these techniques appears to be the principal reason for their transgression beyond mere encoding and into the domain of representations, thus resulting into representation systems and approaches that are increasingly conscious of their limitations, partiality and ambiguities [Brady 1982, 1983].

The decomposition technique proposed in this chapter does not claim to be the only one fit for the analysis and description of architectural plan. It simply fulfills its purpose, that is, to decompose a location into parts so as to facilitate recognition of specific geometric aspects of the location and in particular of aspects which relate to grouping with other locations. The representations used in computer vision are not appropriate (or appropriately calibrated) for the tasks required of CHQP, nor to the kind of objects contained in it and therefore cannot be simply applied to the description of architectural plans. On the other hand, many of their principles and aspects remain relevant to the whole problem and need to be studied more analytically before their integration in CHQP.

## 2.5.2 Quadtrees, medial axis transforms and generalized cones

Davis [1986] distinguishes two basic categories of area-based descriptions for two-dimensional shapes: *partitions*, such as quadtrees and similar hierarchical subdivision schemes, and *coverings*, such as the medial axis transform.

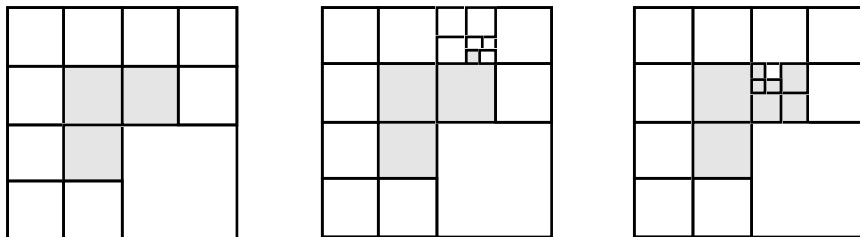
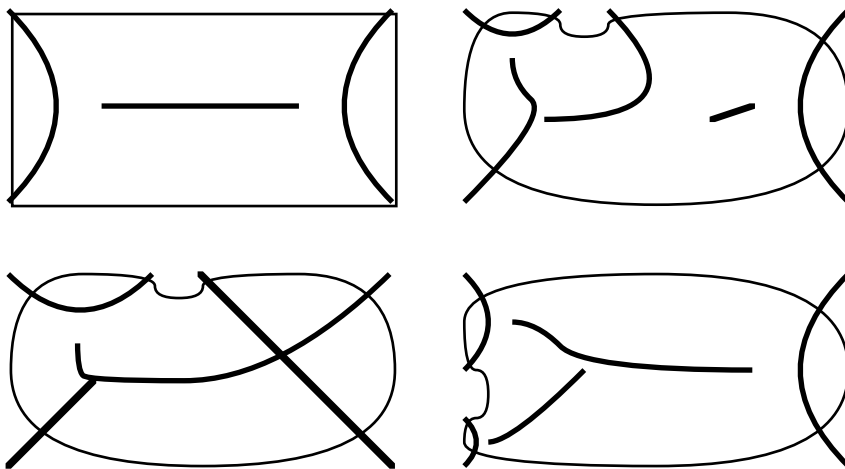


Figure 2-18 Quadtrees

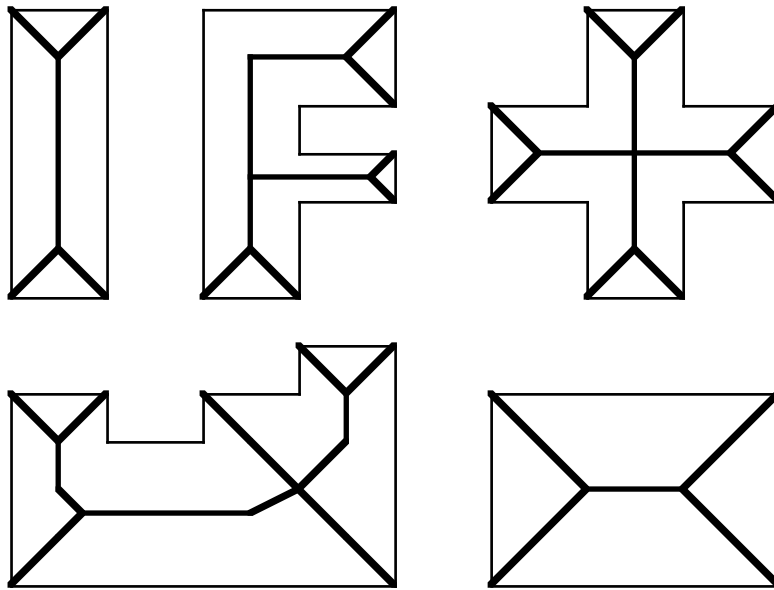
*Quadtrees* are based on recursive subdivisions of an image into quadrants [Samet & Rosenfeld 1980; Samet 1984a, 1984b; Jenkins & Tanimoto 1983]. They primarily aim at reducing computational complexity: instead of heaving to deal with a large number of pixels one deals with the much smaller number of blocks in the quadtree [Davis

1986]. The reduction of elements that have to be manipulated computationally and the straightforward hierarchical structure of quadtree blocks have significant advantages for various manipulation modes. For example, shape descriptions using quadtrees or similar techniques have low sensitivity to local perturbations. Figure 2-18a depicts a quadtree description of an L-shaped area. If a small part is added (Figure 2-18b) or subtracted (Figure 2-18c) the structure of the quadtree does not change significantly, thus facilitating recognition of similarities. Unfortunately this computationally effective and efficient manner of shape decomposition is not always compatible with intuitive interpretations and decompositions of a shape.



**Figure 2-19** *Medial axis transforms, I: extreme sensitivity to relatively insignificant boundary perturbations*  
*practically the same medial axis patterns are produced by notches of any shape and size in the depicted areas of the rectangle*

The *medial axis* (or prairie fire or symmetrical axis) *transform* proposes to be general description of shape. However, this applies more to objects with smoothly curved boundaries and less to rectangular and non-canonical shapes because of the occasionally dramatic response to perceptually insignificant local disturbances in the boundary (Figure 2-19). The medial axis transform was originally proposed by Blum [1967] for describing biological growth and form. This perhaps explains its sensitivity to rather small changes in a shape, as such changes might be indications (first signs) of new growth. Another problem is that in more complex shapes lack of correspondence between the axis and the canonical parts of the shape is quite frequent (Figure 2-20). In short, the medial axis transform is mostly applicable to elongated shapes with gradual changes in orientation and relatively smooth boundaries.



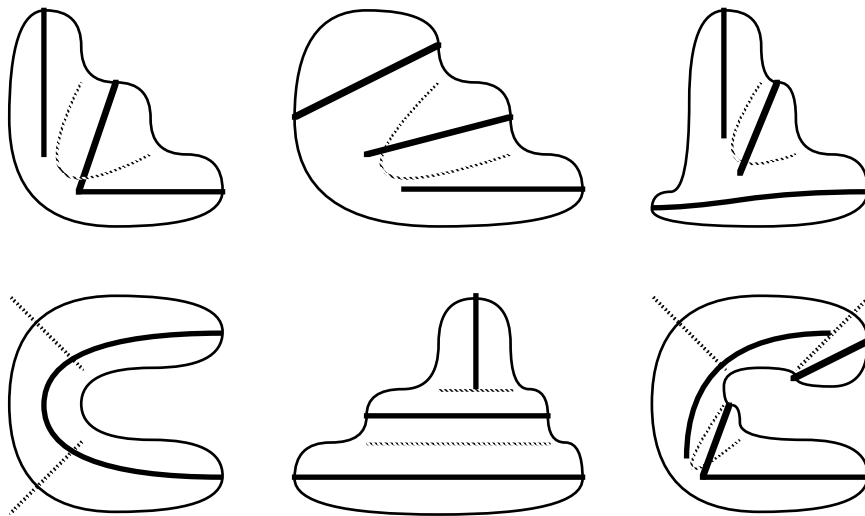
**Figure 2-20** Medial axis transforms, II: correspondence between the medial axis and the boundary of a shape

top row: shapes with good correspondence; bottom row: shapes with poor correspondence: linear silhouettes have in general better correspondence [Pavlidis 1982, p.197]

Similar to the medial axis transform but more concerned with canonical and natural parts is the *generalized cones* (or cylinders) technique. Generalized cones were introduced by Binford [1971] and have been widely discussed and used in computer vision since [e.g. Brooks 1981; Marr 1982]. A generalized cone is described in terms of an axis and a cross section of variable size (and possibly shape) that sweeps along this axis. Generalized cones are often attributed with fewer problems due to boundary perturbations than medial axis transforms. However, as Brady [1983, p.53] has pointed out, *exact* generalized cones suffer very much from such perturbations. In practice, generalized cones are rarely exact. Also the use of several abstraction levels, as in Marr [1982], reduces this oversensitivity. Generalized cones, although essentially developed for the description of three-dimensional objects, are also applied to two-dimensional shapes [Brooks 1981]. In fact, the cones and their axes in a three-dimensional situation are usually identified in two-dimensional projections of their boundaries [Marr & Nishihara 1977; Nevatia & Binford 1977].

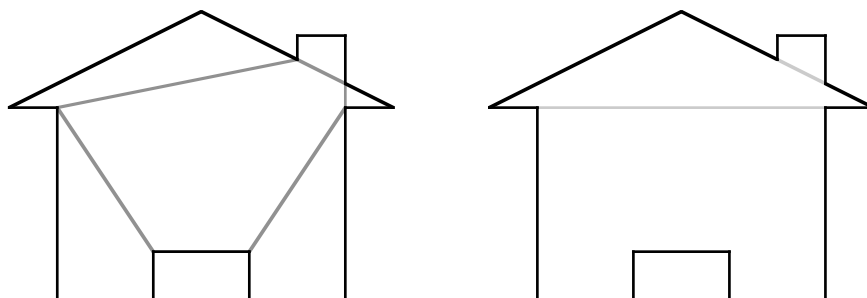
Identification of generalized cones in two-dimensional projections according to Marr & Nishihara [1977] and to Nevatia & Binford [1977] are of particular interest to CHQP, firstly because they involve two-dimensional images and secondly because the identification of the components which comprise an object are the major problem in the use of the generalized cones representation [Brady 1983, p.40]. Once properly identified, the generalized cones of an object offer significant advantages for the structure and manipulation of

the description.



**Figure 2-21** *Recognition of generalized cones, I*  
after Marr & Nishihara [1977]

Marr & Nishihara [1977] identify generalized cones with the help of strong segmentation points in the boundary (occluding contours) of an object. These are typically points of maximum curvature which are interconnected on the basis of heuristic rules to define the outline of each generalized cone. Despite the acknowledged psychological and mathematical significance of maximum curvature points [Attneave 1954; Duda & Hart 1973, p.339], an adaptation of this technique to our case, also influenced by Brady's [1982, 1983] and Pavlidis's [1977, chapter 9] elaborations, does not yield very encouraging results (Figure 2-21). Continuation and repetition in the spatial arrangement and size variations in a shape are not described with adequate accuracy and consistency; segmentation is too sensitive to local perturbations; the overall axis configuration appears to favour hierarchical and canonical structures. According to Kim et al [1987] the reason is that segmentation at deep concave points is not suitable for man-made objects and in particular for rectilinear shapes. In such cases segmentation based on colinearity seems more successful (Figure 2-22).

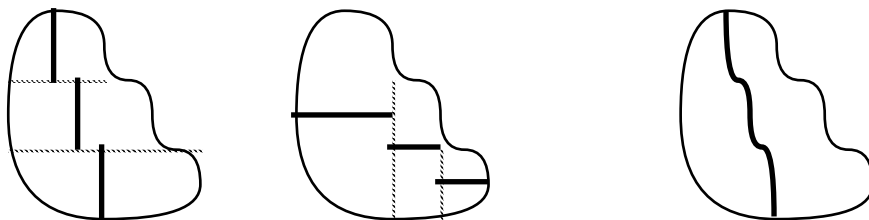


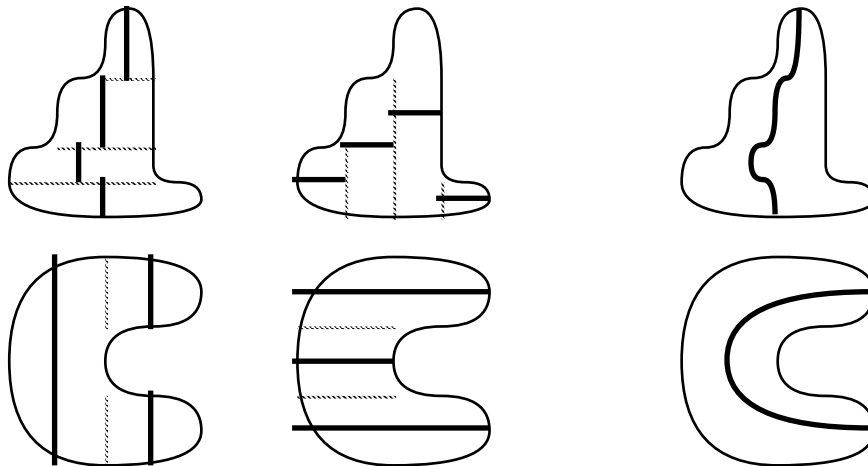
**Figure 2-22 Colinearity in shape decomposition**

*Kim et al [1987] suggest that for man-made objects colinearity yields better results in segmentation (right) than traditional maximal concavity points (left)*

More suited to rectangular shapes is the generalized cone identification technique developed by Nevatia & Binford [1977]. This technique involves the recognition of local cones with straight axes by examining a shape in various directions. The local cones are then extended and/or connected into larger ones which cover the whole object. An adaptation of this technique to the case of rectangular shapes would involve investigation of the two directions, the horizontal and the vertical (Figure 2-23). This technique is more successful in capturing continuation and repetition but still fails to combine these with peripheral conditions which constrain its relations with neighbouring shapes. The description of a location by its central axis does not say much about its boundary where it contacts its contiguous locations in manners prescribed by, for instance, construction or site topography constraints.

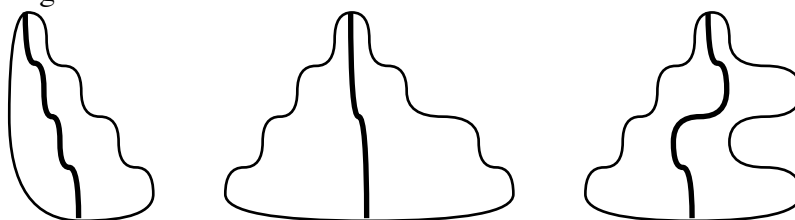
Generalized cones are well suited to describing organic forms but require reconsideration if applied to orthogonal shapes [Brady 1982, 1983; Pavlidis 1977, chapter 9]. Generalized cones are sensitive to orientation (aspect ratios) and this leads to a strong preference for elongated and canonical forms. Regular changes in a shape are accentuated, as for instance in Figure 2-24a where the straightness of the left and bottom sides are ignored by the central axis. Generalized cones are more suited to describing forms characterized by axial bilateral symmetry. For example, Figure 2-24b is better described by generalized cones than Figure 2-24c. Although generalized cones capture the significance of these axes (especially when the axes are related to hierarchical subdivision, as in the human body) they do not fully describe the boundary of an object but give instead a kind of summation of the two opposite sides along the medial axis.





**Figure 2-23** *Recognition of generalized cones, II*  
 after Nevatia & Binford [1977]; the left column shows the results of investigations for local cones with vertical axes, the central column shows the results of investigation for local cones with horizontal axes, and the right column the combination of the two on the basis of the following criteria: maximal total length of axis, minimal number of perturbations on the axis, and minimal shift of direction in axis perturbations

In chap the significance of the boundary is undeniable because it is there that the contextual (extrinsic) constraints (for instance those deriving from the load bearing structure) are most clearly expressed and also because boundaries determine to a large degree the spatial and geometric relations between neighbouring locations and subsequently constrain their grouping. For example, alignment of Figure 2-24a to other locations with respect to its left or bottom side is more significant than alignment with respect to some part of its top or right side because of the fragmentation of the latter two sides.



**Figure 2-24** *The applicability of generalized cones to the recognition of two-dimensional orthogonal shapes*  
 left (a): the axis describes the jagged right side of the shape but appears not to account for the smooth left side; middle (b): generalized cones are better suited for describing shapes characterized by axial symmetry; left (c): asymmetrical shapes are not properly described even though the sides which influence the position and shape of the axis are both characterized by regular patterns

This criticism does not aim at disproving the validity or the utility of generalized cones but simply at pointing out the aspects of the generalized cones representation that should be modified for their application in chap. Whether the procedures proposed for their identification by Marr & Nishihara [1977] and Nevatia & Binford [1977], and perhaps even the elaborations by Brady [1982, 1983] and Pavlidis [1977, chapter 9], should be restricted to describing animal

or similarly articulated forms (i.e., with elongated limbs) is beyond the scope of the dissertation. What is significant for the dissertation is (a) definition of criteria for the selection of the appropriate techniques for the description of locations in architectural plans and (b) investigation of the precise nature of changes required for the adaptation of such techniques and representations to their new context.

### 2.5.3 Decomposition into rectangular slices

The generalized cones representation appears to be the better choice among the techniques discussed in the subsection 2.5.2 for two basic reasons. The first is that generalized cones offer the means for recognizing the articulation of a shape by making explicit their subdivision into parts and the second that their extensive and intensive use in computer vision provides a variety of tools that can be adapted for the purposes of *chap*.

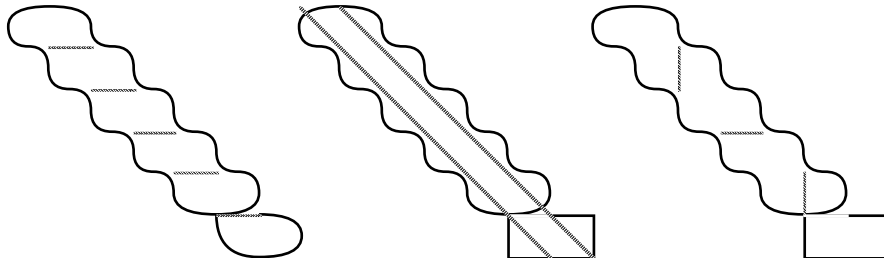
Of the two techniques for the identification of generalized cones in an image examined in 2.5.2, the one by Nevatia & Binford [1977] performed better within the totally orthogonal environment of *chap*. This suggests that a primary merit of the decomposition technique that should be adopted is correspondence and consistency with the orthogonality constraints of *chap*.

The major problem with the applicability of both the technique by Nevatia & Binford [1977] and the one by Marr & Nishihara [1977] to *chap* is the representation of the configuration of generalized cones in an image by the configuration of their central axes. As we have seen in 2.5.2, this is mainly due to the differences between the images that normally concern computer vision and the totally orthogonal architectural plans of *chap*.

Computer vision systems are generally concerned with images of objects such as animal limbs which are well suited to axial descriptions. Axial descriptions capture the articulation of their skeletal system and subsequently can represent the various postures of the limbs through the spatial relations of their axes, thus forming a global hierarchical model comprising, on the higher levels, of the axes and their relative positions.

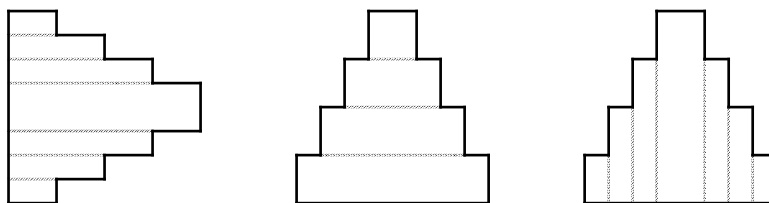
For locations of an architectural plan similar abstraction might be useful for general shape categorization. However, even if successful, such categorization would be insufficient as a description of shape and spatial articulation. Locations consist of clearly or vaguely distinct parts; whatever the approach one adopts with respect to the subdivision of locations (usually there are quite a few

equally acceptable alternatives), there are distinct and undeniable patterns of spatial arrangement common to most of the alternatives. For example, regardless of the subdivision one might accept for Figure 2-25, there is a distinct pattern of repetition and translation of identical or parts with similar features.



**Figure 2-25** *Alternative decompositions of a shape*

In order to capture these patterns and provide links for the correlations of the articulation of one location with that of its neighbours the proposed technique of shape decomposition shifts attention from the central areas of a shape to its limits. In generalized cones attention has focused on the central areas and in particular to central axes. Both the medial axis transform and generalized cones are concerned for the ‘spine’ of a shape because it usually corresponds to a natural or canonical spine. In architectural plans such ‘spines’ are mostly not found at the centre of locations but rather at the boundary. For example, axes of load bearing elements which constrain the shape, size and position of a location are usually encountered at its boundary. An examination of the constraints apparent in each location can thus reveal aspects of the load bearing structure of the whole plan which also relate to the grouping of locations, while a medial axis can reveal directly only the general orientation and arrangement of the locations in the plan. Although this is a major revision of the generalized cones representation, it should be noted that salience of the boundary over the area that is not adjacent to the boundary holds not only for architectural plans but also for an issue that has attracted tremendous interest in computer vision, texture discrimination [Muller 1986].

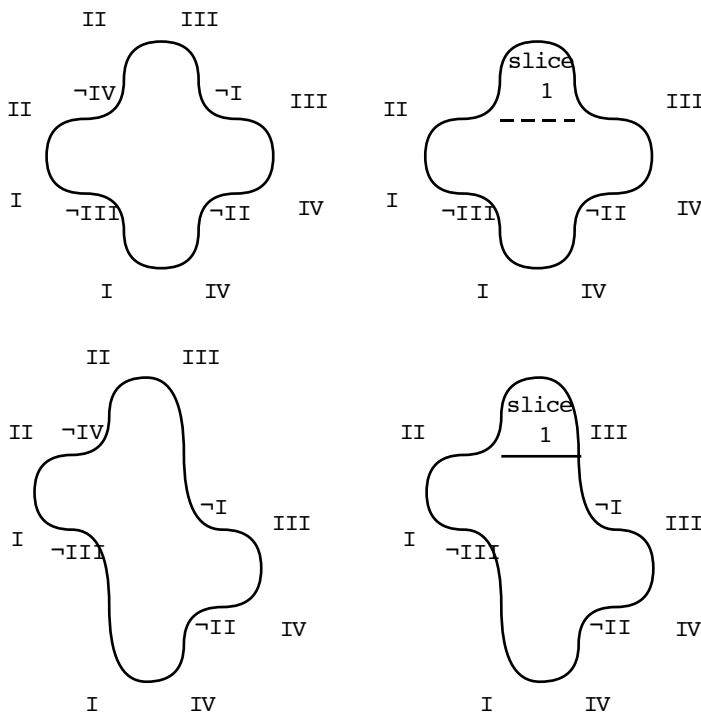


**Figure 2-26** *Horizontal and vertical slices*

*decomposition into horizontal rectangular slices facilitates recognition of vertically oriented relationships, such as vertical alignment fronts (left: the left edge); although horizontal slices can be used for the recognition of the same relationships in the horizontal direction (middle: the bottom edge is an alignment front), equivalence and consistency are enhanced when the relationships are detected and explained through decomposition into vertical slices (right)*

The proposed technique of shape recognition is based on the assumption that, if segmentation of an orthogonal shape into elementary rectangular components is performed in a consistent and unambiguous manner, the resulting set of components can be used to extract general structural characteristics which are specific of each shape type. Since the purpose of this segmentation is to facilitate detection of structural shape constraints at the boundaries of each location, it was assumed that all resulting component rectangles should be directly related to the boundary. Subsequently, hierarchical decomposition techniques like quadtrees were disqualified.

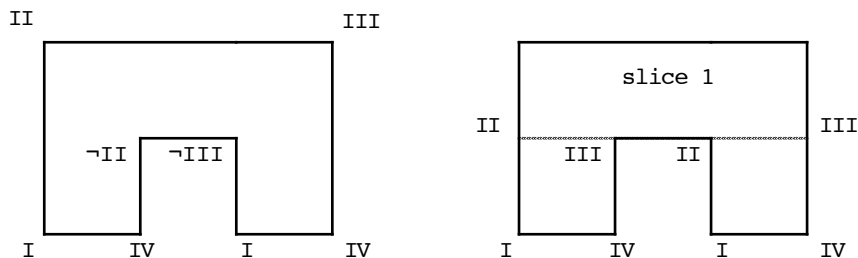
One technique that seems to provide acceptable results is *decomposition into rectangular slices*. Following the constraints of the row-major (lexicographic) order of arrays in Common Lisp, the programming language used for the development of *chap* [Steele 1984, p.28], a location is generally decomposed into *horizontal* slices which facilitate recognition of slice relationships with vertical to diagonal orientation. Relationships with horizontal orientation can also be recognized through decomposition into horizontal rectangular slices but for clarity and consistency's sake decomposition into vertical slices is also applied (Figure 2-26).



**Figure 2-27** *Decomposition into horizontal slices: the procedure used in chap*

*top left (a) and bottom left (c): the initial vertex configurations of two shapes; top right (b) and bottom right (d): the modified vertex configurations after the recognition of the first slice*

Decomposition into horizontal rectangular slices is performed by a procedure which utilizes the constraints integrated into the vertex typology of Figure 2-10. The procedure starts by identifying the top left vertex (in row-major order) of type II and proceeds to identify the two vertices connected to it. Then the fourth vertex of the rectangle is deduced from the coordinates of these two. If this fourth vertex does not correspond to a vertex of the location it is added to the vertices of the location (*virtual vertex*). The four vertices define a rectangular slice which is segmented from the location. This segmentation incorporates modification of the vertex configuration of the location: all vertices of the slice are removed if they correspond to vertices of the location (Figure 2-27b), while virtual vertices are added with type III (Figure 2-27c). The next slice is then recognized on the basis of the modified vertex configuration and so on until no location vertices are left.



**Figure 2-28** More elaborate cases in decomposition into slices  
left (a): the initial vertex configuration; right (b): the modified vertex configuration after the recognition of the first slice

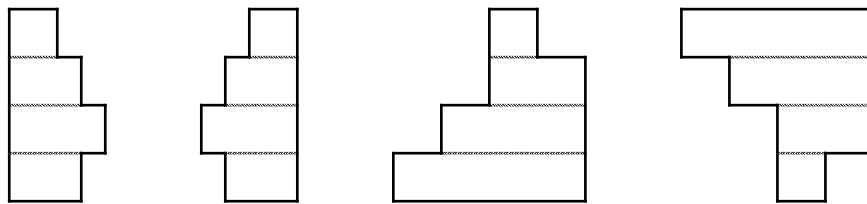
An elaboration to the procedure concerns cases where the rectangle defined by the type II vertex and the two vertices connected to it might include in its area other vertices (Figure 2-28a). In such cases the row of the top left of the latter vertices determines the bottom edge of the slice. The vertex configuration is modified in a different way than the one described above, as the slice usually involves two virtual vertices. Also the type of the two or more other vertices on the bottom edge of the slice is modified (Figure 2-28b).

#### 2.5.4 Relationships between the components of a location

Recognition of the shape of a location is based on the geometric / topologic relationships between the components produced by the decomposition of a location and the resulting arrangements (groups) of components. In the initial version of CHAP these relationships have been reduced to three basic ones which can account for any type of

component arrangement: alignment, transposition and alternation.

*Alignment* means that a specific edge of the components forms a continuous straight line segment, called the *alignment front* of the resulting group (Figure 2-29). The criteria according to which a set of components constitutes an alignment group are, e.g. in the case of a left-side alignment front, that the top left vertices of all locations are on the same column and that the bottom left vertex of each component (except for the last) is adjacent to the top left vertex of the next. The order of the components is the order in which they were recognized during decomposition, i.e., is defined by the row-major order of their top left vertices. Adjacency assumes here a more rigorous sense than in everyday language: two vertices are adjacent only if they either belong to the same column and are separated by one row, or belong to the same row and are separated by one column.

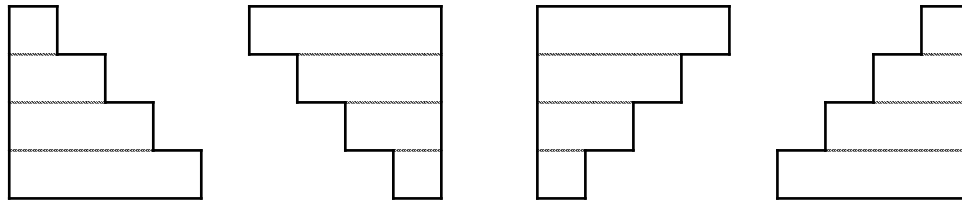


**Figure 2-29** Examples of alignment groups  
(from left to right) left-side, right-side, bottom, and top alignment front

Horizontal alignment fronts can be detected among horizontal slices by checking the relative size and position of the bottom edge of one component against those of the top edge of the next component. Care is taken to ensure that the alignment front corresponds to an edge or part of an edge of the location.

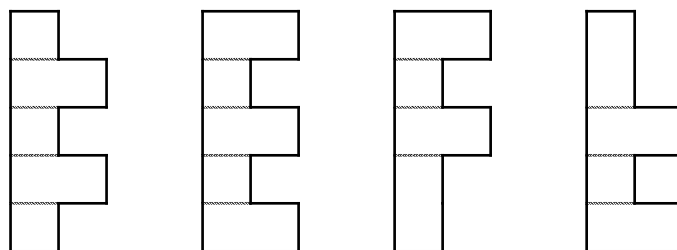
A *transposition* group is one where the position a specific edge of the components changes in a steady manner (Figure 2-30). By 'steady manner' it is meant that the specific edge of one component can be obtained by a translation and/or scaling of the corresponding edge of the previous component. Both translation and scaling can be quantitatively variable in each instance; only the direction and sense of the translation need be roughly the same, i.e., the translation must be produced by the addition of a non-zero quantity which for each group can be variable but must remain either positive or negative. Negative quantities characterize *ascending* transposition patterns, while in the case of positive quantities we have *descending* transposition. The criteria for recognizing that a set of components is characterized

by, for instance, descending transposition on the right side are that the column of the top right vertex of the first component is of lower order than the same vertex of the second component, which in turn is of lower order than the same vertex of the third component, and so on.



**Figure 2-30** Examples of transposition groups  
(from left to right) right-side descending transposition, left-side descending transposition, right-side ascending transposition, and left-side ascending transposition

The definition of *alternation* is similar to that of transposition, with the difference that translation is produced by the alternating addition of positive and negative quantities (Figure 2-31). There are different kinds of alternation, depending upon the side of the location and upon whether the first and last translation are produced by negative or positive quantities, that is, upon the pattern of projections and recesses. The criteria for recognizing a set of components as a group characterized by, for instance,  $a-b-...-a$  alternation on the right side are that the column of the top right vertex of the first component is of lower order than the same vertex of the second component, which in turn is of higher order than the same vertex of the third component, and so on, while the column of the top right vertex of the last component is of lower order than the same vertex of the previous component.



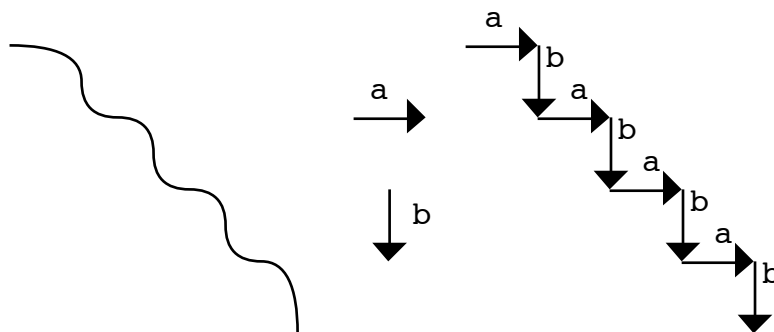
**Figure 2-31** Examples of alternation groups (from left to right): right-side  $a-b-...-a$  alternation, right-side  $b-a-...-b$  alternation, right-side  $b-a-...-a$  alternation, and right-side  $a-b-...-b$  alternation

## 2.5.5 Recognition and spatial arrangement of component groups

So far we have discussed conditions of alignment,

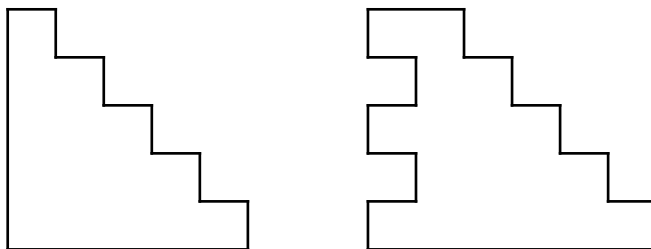
transposition and alternation among components of a shape in terms of edges, that is, in a manner reminiscent of traditional structural pattern recognition techniques (Figure 2-32). For example, we might examine each side of a shape to determine which patterns occur and then produce somehow a unified description of the shape. However, the applicability of this technique is restricted to rather simple shape types (mainly linear arrangements, i.e., with two primary sides) because correlating the patterns of one side with the others can create unacceptable ambiguities in the overall interpretation of the shape. Moreover, such techniques bring us back to line following, with all disadvantages discussed in previous sections.

The most significant of these disadvantages is that knowledge of the perturbations of a location does not constitute direct knowledge of more general spatial patterns nor of the general shape of the location. In order to recognize these one must develop elaborate correlations of the metric properties of edge patterns (Figure 2-33) but the complexity and tenacity of such correlations rules out efficiency in the case of more complex shapes (Figure 2-36).



**Figure 2-32** Recognition and coding of a two-dimensional edge pattern

*the pattern is described in terms of the primitives a and b as the string abababab (after Gonzalez & Thomason [1978])*



**Figure 2-33** Description of shapes in terms of its edge patterns

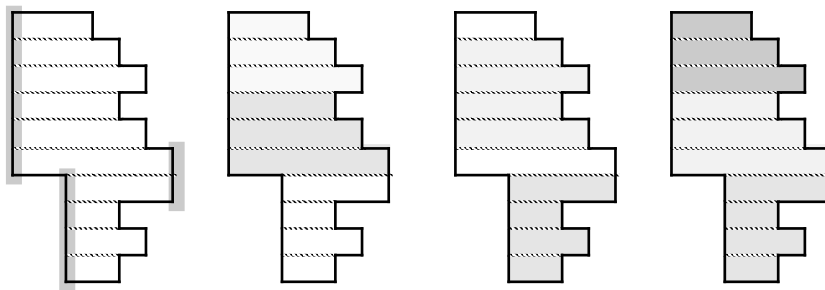
*left: a shape that is characterized by alignment (left side,) transposition descending (right and top) and alignment; right: a shape that is characterized by alternation b-a-...-b (left side), transposition descending (right and top) and alignment (bottom)*

By considering the same relationships between spatial components of a location, such as the rectangular slices used in chap, it is possible not only to correlate different edge

patterns but also, in principle at least, to identify the parts of a location in the form of groups of components and thus derive accurate descriptions of the shape of a location as well as of its spatial articulation.

The grouping procedures for location components in *chop* are explained below through a number of examples. Recognition of groups in these examples is described primarily in terms of the algorithmic structure of the procedures and their coordination. Implementation details are left out, as also are secondary algorithmic operations which do not affect recognition of groups in all cases. These include most cases of investigation for horizontally oriented grouping relationships, as in the examples investigation for vertically oriented relationships generally exhausts grouping possibilities.

The first example concerns a rather simple case which nevertheless involves resolution of the most frequent ambiguities and illustrates the priority rules employed in *chop* (Figure 2-34). Identification of groups commences with alignment, the strongest of grouping relationships. In the example three alignment fronts are observed (Figure 2-34a). Of these one is dismissed because the related alignment group consists of only two components. Groups with less than three components are not taken into consideration except when one or both components cannot belong to other groups.



**Figure 2-34** *Recognition of component groups, 1*

(from left to right) (a): alignment fronts (denoted by thick gray lines); (b): transposition groups; (c): alternation groups; (d): resulting discrete groups

The second step in group identification is the investigation of transposition groups. In the example two acceptable transposition groups are identified (Figure 2-34b). Two others (not shown) are dismissed because they consist of only two components. Although alternation is stronger than transposition in *chop*, precedence is given to the investigation of transposition because alternation may occur as a byproduct of other forms of spatial arrangement, for instance where two transposition groups meet, as in the present example.

The investigation of alternation returns two acceptable groups in the example (Figure 2-36c). Alternation cases not accepted as groups are not shown. Of the not accepted

alternation groups one coincides with an alignment group and is disjoint with the two transposition groups identified previously. The other does not coincide with any other group but partially overlaps with three of them (the top alignment group and both transposition groups).

The overlapping and multiplicity of groups raises the issue of the relative acceptability of alternative groupings. In principle there is no reason why one should not accept multiple interpretations of component grouping, as there are quite a few ambiguous cases in perception (Wittgenstein's duck-rabbit, Necker's cube, Rubin's vase and all illusory contour figures). However, multiple interpretations should be true alternatives and not accidental misinterpretations.

At this stage of *chap* I am not inclined to consider acceptability of alternative group configurations in terms of information theoretic or similar economy [Attneave 1954; Hochberg & McAlister 1953; Garner & Clement 1963; Garner 1974; Palmer 1982, 1983] because, even though such issues might contribute to aesthetic appreciation [Prak 1977] (with all reservations raised by the criticism of the applicability of information notions in discussions of the meaning of a representation [Black 1972, pp.104–110]), there is no evidence that they govern the structure of architectural design, at least not directly.

Rock [1982] suggests an alternative approach by considering perception as a form of problem solving where preference for a solution is based on the most complete explanation of the stimulus in terms of a coherent set of common causes. In our case, locations in architectural plans are products of human design. Therefore, we can assume that such common causes (i.e., formative design techniques and basic principles) do exist in the configuration of the shape of a location and that this shape is determined by these causes.

On the basis of this approach a number of criteria for the acceptability of group configurations have been set up. One of the most significant is the *coherence* of a group. Coherence is measured by the number and type of relationships between the components of a group. For example, an alignment group is more coherent than an alternation group because alignment is seldom if ever accidental. In turn, a group characterized by two relationships, such as alignment on one side and transposition on another, is more coherent than a group characterized by one relationship only. The other primary criterion concerns the relations between group configurations: if one configuration can be produced from another by the union of two or more groups, then only the most analytic configuration is accepted, since relationships between groups and their union into more comprehensive structures are investigated at the following stage. Both criteria, as well as ancillary ones not mentioned here, are not

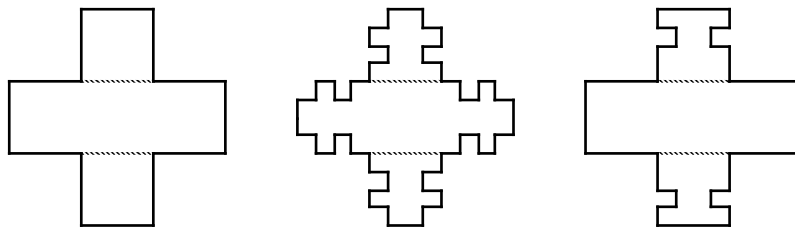
described in detailed because in their current form they constitute only first approximations of a future definite control structure.

In the example of Figure 3-34 we have two alternative configurations of disjoint groups: one of two groups and two solitary locations (one on top and the other between the two groups) and one of three groups. In principle both configurations are equally acceptable. The latter might denote coherence and economy in the use of formative techniques in architectural design, while in the former the two solitary components, one at the end of the location and one between the two groups, can be explained respectively by the influence of some extrinsic (contextual) constraints and as a transition between the two groups. This means that either component cannot be grouped with others in the specific group configuration. However, in Figure 3-34 differentiation of the two solitary components is destroyed by the fact that both are aligned with one of the two groups, thus forming a definite larger group which cannot be ignored. Therefore, the former configuration actually comprises two groups only and no solitary components. As such, it can be produced by the other configuration, by the union of two groups. Subsequently, of the two configurations in the example only the one is acceptable, that of three groups (Figure 3-34d).

The final stage in group recognition involves correlation and coordination of the discrete groups that have been identified. This stage corresponds to the classification of the overall articulation into one of a finite number of classes found in most shape recognition systems. Each group is treated as if it were a single component and conditions of alignment, transposition and alternation are again investigated. The change of scale involves a shift from the strictly metric criteria used in the investigation of the same relations between components towards more topologic aspects. In the example the two first groups from the top form on the left a unified alignment front, while on the right they are both characterized by the same type of transposition. Such homogeneity (probably the result of a conscious and consistent combination of the same design constraints if the shape corresponded to a location) transforms the initial tripartite subdivision of the whole into bipartite: the resulting description is one of a linear arrangement with two subdivisions, one of two aligned and similar groups, and one of a single group with a displaced alignment front.

Descriptions on this level of abstraction allow us to ignore local perturbations and concentrate on the structure of the shape and also distinguish between such accidental or ad hoc perturbations and the particular articulation of each part of a location. For example, all cross shapes depicted in Figure 2-35 are similarly described as tripartite a-b-...-a alternation

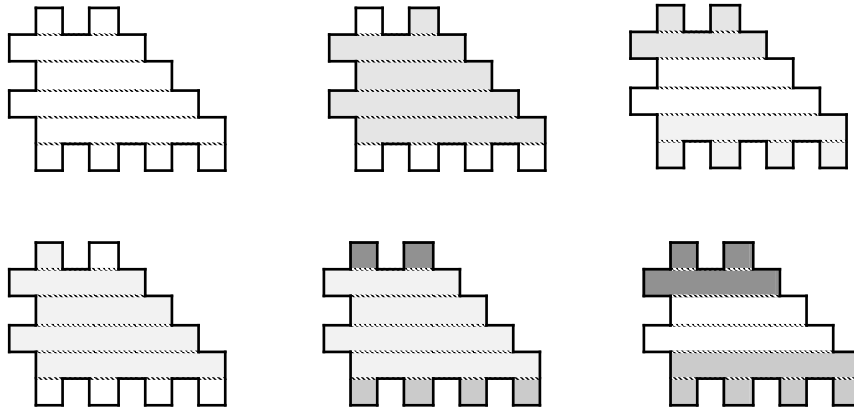
configurations, although in each case the subdivisions are different in nature and composition. In the one shape they are three single components, in the other three groups of components and in the last one two groups and one component. Conformity with canonical shape categorization is not among the goals of CHQP because it often poses insurmountable ambiguity problems. On the other hand, compatibility with the intuitive shape interpretations which underlie canonical shape categorization is significant because we are dealing with consciously designed entities which follow the specific principles of shape articulation. In that respect, similarity of description between intuitively similar shapes is very useful in the analysis of possible formative techniques used in architectural design and also for the development of efficient user interfaces.



**Figure 2-35** *Cross shapes described by similar structures  
all three shapes are described as a tripartite a-b-...-a alternation,  
although subdivisions are different in each case*

While most location shapes are less complex than Figure 2-34, more complex shapes which often defy direct interpretation on the basis of line following are perhaps more enlightening as to the potential of the proposed technique of shape description. One such case is depicted in Figure 2-36. This example has no alignment fronts and only one transposition group (Figure 2-36b). The investigation of alternation reveals three partially overlapping groups which cover the entire shape (Figure 2-36c, d).

Correlation and coordination of these groups reveals that there is only one group which is characterized by two relationships, the one formed by the intersection of the transposition group with the largest of the alternation groups. What remains is components which are not directly related to each other by alignment, transposition or alternation (Figure 2-36e) but still form two distinct groups and appear to follow a certain pattern. In fact, they form subsets of two alternation groups (Figure 2-36c). The removal of a single component from either group destroyed continuity and also the formal status of the group. The removal of the two components is unavoidable because they also belong to a stronger group (one with two relationships between their components) and cannot be reversed so as to conserve the unity of the two weaker groups.



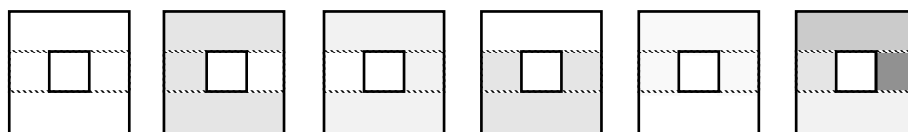
**Figure 2-36** Recognition of component groups, II

top left (a): the shape and its components; top middle (b): transposition groups; top right (c) and bottom left (d): alternation groups; bottom middle (e): resulting discrete groups; bottom right (f): completed virtual groups

On the other hand, we cannot ignore that the components are related, even if indirectly, nor that their unity is destroyed by an artificial decomposition scheme and an equally artificial grouping into disjoint groups — artificial in the sense that in reality the subdivision of a continuous parcel of space into precise and discrete particles always involves a certain degree of fuzziness. The problem is tackled by making allowance for *virtual groups*, that is, groups which are not complete but can be completed with the addition of one or two components that can be taken from another group without destroying its structure and continuity.

In the example of Figure 3-36 the solitary components form two virtual groups with the addition of one component from the already established group (Figure 2-36f). Virtual groups are considered without these additional components except under certain special circumstances, for instance when the precise contour of the alternation front is required.

Virtual groups are particularly useful in cases where component groups of similar coherence overlap to a degree such that does not allow for dismissal or reduction to virtual just for some of them. One such case is locations which enclose a hole (atrium) or another location (Figure 2-37). Here there are four overlapping groups of equal coherence. Each group consists of three components and is characterized by alignment on the one side and alternation on the other. In this case no preference can be established for or against any of these groups and therefore each component constitutes a virtual group.



**Figure 2-37** Recognition of component groups, III

(from left to right) (a): the shape and its components; (b, c, d, e): the four alignment/alternation groups; (f): the resulting four virtual groups

The process of grouping components of a location does not result in a verbal or typologic characterization (label) of its shape. This means that two similar shapes might be described slightly differently (not very much though) and that there is no shape type categorization in CHAP. Instead, shape is described in terms of patterns which do not reflect directly the terminologies and shape typologies used in everyday language and in architecture. Although there is no reason why such typologies and terminologies cannot or should not be superimposed on the grouping process described in this subsection, such augmentation is not required for CHAP at this stage of its development.

## 2.6 Conclusions and extensions

Recognition of locations in CHAP consists of two main components, identification of the vertices of a location and shape recognition through decomposition into horizontal slices. The first component has strong relations with human perceptual mechanisms and with successful computer vision systems. These, together with the effectiveness and reliability of the proposed technique of vertex identification, justify the adoption of the vertex-based approach to recognition in CHAP. The second component, shape recognition through decomposition into horizontal slices, supports abstraction of shape description in a manner comparable to any computer vision system but a more comprehensive justification for the adoption of the decomposition approach to shape recognition is provided by the correspondence of this component to the grouping of locations (cfr. Chapter 3). Therefore, the present section avoids all aspects of the justification of recognition of locations in the architectural plans of CHAP and concentrates on the extensions of the proposed techniques.

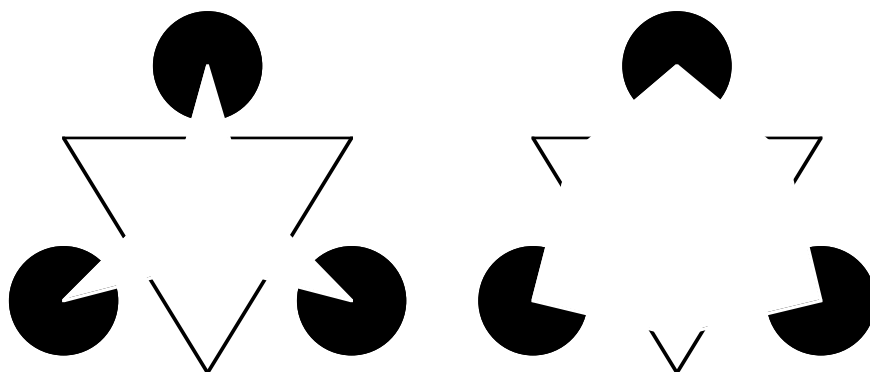
### 2.6.1 Beyond orthogonality

The proposed system has been developed for a totally orthogonal environment but can be easily extended beyond orthogonality. Such extension does not require changes in the structure of the system but only augmentation of its constraints, as suggested by Walters [1986] who observed that the same preferential enhancement hierarchy (Figure 2-9) applies to both rectilinear and curvilinear line drawings and also by Fleck's [1986] local rotational symmetries extension of Brady's [1982, 1983] smoothed local

symmetries representation. The augmentation affects three major groups: the typology of vertices, the connectivity constraints between vertex types, and decomposition criteria.

Non-orthogonal architectural plans require that the system can recognize additional vertex types to the eight basic ones (Figure 2-10). An appropriately augmented typology may comprise sixteen to sixty four types, depending upon the range of shapes encountered in the application domain, the resolution of the system and the degree of line normalization in preprocessing. The number of types can be further reduced by a hierarchical organization of the types through the use of different resolution levels, as in Marr [1982].

This hierarchy is also useful in organizing connectivity constraints into a computationally efficient structure. As long as we are dealing with rectilinear shapes only, connectivity poses no serious problem. Curvilinear shapes are more challenging because it is not always possible to predict the shape of a curvilinear edge from a small part at either of its vertex. In these cases line following seems at first sight a reasonable although computationally expensive solution, especially in the case of complex and irregular curves.

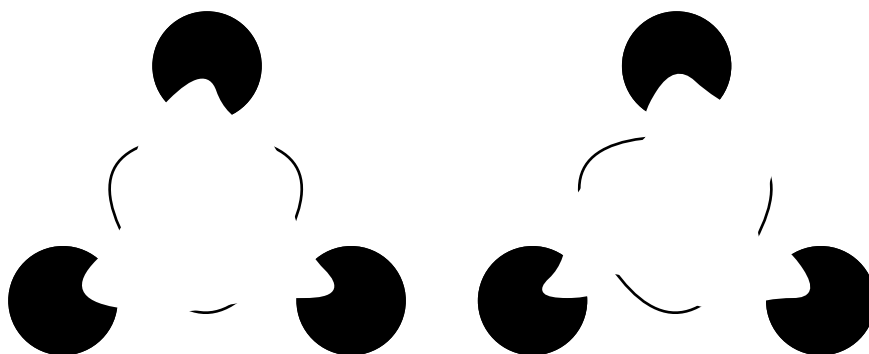


**Figure 2-38** *Curvilinear illusory contour figures, I*  
left (a): with concave angles; right (b): with convex edges

However, line following is not necessary, as we can see from curvilinear illusory contour figures (Figure 2-38). Recognition of curvilinear illusory contours can be achieved by forming expectations for the direction of connected vertices as in *CHAP*, with the addition of direction tolerances. Such tolerances would specify that a connected vertex should lie within an angle from the other vertex rather than on a specific row or column. For example, we could specify such tolerances on the basis of extreme examples, such as those in in Figure 2-38, and apply them in the recognition of more complex illusory contour figures, like Figure 2-39a.

An interesting observation is that the perceived form of the illusory contours is not directly produced by

linking all the supposed occlusion points but rather is constrained by them. The illusory contours in Figure 2-39a are perceived as complex curves which link each pairs of vertices. The curves lie *within* rather than *on* the supposed occlusion points of the boundary of the occluded triangle. The reason for these discrepancies appears to be a preference for smooth changes in the orientation of the edges which link the three vertices (the good continuation criterion of Gestalt theory). The preference for smoothness is even more evident in Figure 2-39b, where the smooth illusory contour figure defined by the three vertices is so severely disrupted by the lack of correspondence with the perceived occlusion of the triangle boundary that perception of the illusory contours is quite weak. As a matter of fact, both Figure 2-39a and 2-39b were actually created by drawing a rectilinear irregular figure on top of three black disks and a triangle. In both cases the vertices of the occluding figures lie almost exactly on the perceived occlusion points. The significance of this preference for smoothness is an aspect that should concern extensions of the proposed technique but has no effect in the orthogonal environment of chap.



**Figure 2-39** *Curvilinear illusory contour figures, II*  
left (a): an illusory contour figure defined primarily by its vertices  
(note that the figure appears to be within rather than on the  
occlusion points of the boundary of the occluded triangle); right  
(b): the stimulus does not support perception of an illusory contour  
figure because of lack of smooth transition from one supposed  
occlusion point to the next

## 2.6.2 Addressing shape variety in architectural plans

Form description in CAAD is almost exclusively based on constructive geometry, i.e., set-theoretic operations on a limited number of primitive geometric forms, generally

regular shapes and volumes. The practical advantages of keeping the number of primitive forms small are obvious and do not pose conceptual discrepancies with current conceptions of architectural design. On the other hand, the set-theoretic operations are not sufficient for describing accurately the relations between the components of a shape or volume: the overall spatial articulation of the components is always implicit in the spatial properties of each component and in the relations of union, intersection or difference between each pair of components. As a result, the overall structure of each design entity is reduced to a simple aggregate of component properties.

The types of spatial arrangement examined by chap in the recognition of locations, alignment, transposition and alternation, form the basis for an additional level of relations between components, specifically aimed at describing the structure of the whole in abstract terms, thus facilitating its manipulation and categorization. This description is instrumental for the automated identification and manipulation of descriptions of design entities in meaningful manners which correspond to the properties of these entities and not to the properties of the implementation mechanisms used for the representation.

A currently dominant alternative in CAAD and also in Artificial Intelligence is the abstraction of aggregates of primitive components into higher-level entities on the basis of rule-based whole-part hierarchies. These hierarchies arguably provide efficient resolution of abstraction-related problems in some simple tasks (i.e., when the rule base is rather compact and unambiguously structured) or in conjunction with strict input structures (i.e., when the user analytically and strenuously differentiates aspects and attributes of the input descriptions or when the computer can perform such operations automatically). However, they are often unable to tackle large and complicated problems on the basis of the loose and fragmentary data typically encountered in the design process.

The effectiveness and adequacy of the technique proposed in this chapter should not be measured by the comprehensiveness of the arrangement types used. I have no doubt that different perspectives and concern for different aspects of building geometry might involve other types of spatial arrangement. The procedures outlined in this chapter should be viewed as the basic module of a general purpose system of location recognition in architectural plans. This module can be readily augmented and elaborated with several additional criteria. These include the *aspect ratio* of components, which effectively corresponds to their orientation, various kinds of *symmetry* (thus integrating all essential geometric transformations) and complex types of spatial arrangement which can be analysed into sequences of

instances of the three basic types.

### 2.6.3 Decomposition

Decomposition into elementary primitives has been one of the few common grounds in computer vision. Some decomposition techniques have been highly praised for the correspondence of their output to natural (canonical) subdivisions. Others are appreciated for their computational advantages. Overall, however, the acceptance of decomposition depends more upon the general perspective adopted with respect to human perception than to the particular merits of each technique. Perspectives influenced by Gestalt theory tend to relegate the significance of decomposition and parts and stress instead the types and criteria of spatial arrangement. Others have been less reluctant to reduce an image to a features vector or a relational graph of parts. In Artificial Intelligence the addition of sophisticated inference mechanisms represents an attempt to relate individual features to global aspects via heuristic rules.

Middle-of-the-road approaches are, not surprisingly, more attractive. In particular, the use of different resolution levels by Marr [1982] performs well with respect to the production of a hierarchy of abstraction levels and, for canonical views at least, reduces the problem of part-to-whole relationships to an unambiguous hierarchy of relational structures of parts. Marr's approach is instrumental in understanding how a whole may be synthesized out of its parts, but does not provide clues on how the reverse, the analysis of a whole to its parts and the relationships between the parts, might be achieved. As a result, the approach is vulnerable with respect to low-level problems [Watt 1987; Mowforth et al 1987] to a degree inconceivable for e.g. the Gestalt principles of organization. Criticism extends to the point that "results obtained by workers like Marr ... are considered at odds with human interpretation where the human sees edges that these operators have omitted to label or alternatively labels have been provided where the human does not see any edge." It is suggested "that these approaches fail because they make both strong and unrealistic assumptions about intensity changes" [Mowforth et al 1987, p.177].

Recently there has been a notable increase in the significance of linking perceptual decomposition to decomposition in computer vision [Biederman 1985, 1987]. This tendency may be significant for computer representations of visual images, as we need more information on how decomposition is performed and how it affects form recognition. For example, Attneave [1954] has

shown the significance of maximum curvature points in drawings comprising edges which link maximum curvature points. No-one has any difficulty in perceiving the drawing as one of a cat. However, as Green & Curtis [1966] have pointed out, no-one perceives the cat in a drawing consisting of only these points; the direction and order of connecting edges is necessary. Vertices contain such information but it is still not clear to me how they might relate to each other in image segmentation.

One thing seems rather clear: it is certainly worth investigating the hypothesis that perception is based on symbolic, abstract descriptions based on decomposition and grouping and not on precise, analytic enumeration of every minute detail in a shape [Rock 1983, chapters 3 & 4; Marr 1982, chapters 1 & 2]. Decomposition appears to be present from early on. Baddeley [1986] presents good evidence for a short-term (or working) memory where the auditory and visuospatial stores are clearly distinguishable. In the latter, the visuospatial sketchpad (VSSP), where images are maintained and manipulated, decomposition appears to play a significant role. As Frick [1989] has observed, visual confusion errors (that is, the tendency of incorrectly reported items to contain a part of the correct item) suggest that at least alphanumeric characters in the VSSP have parts which degrade independently. Prototypical representation of alphanumeric characters in the VSSP is also dismissed, as there were errors which showed influence of the font of the character.

There can be many objections to the non-canonical and seemingly arbitrary decomposition of shapes in the location recognition process proposed in this chapter. However, there are a few perfectly good reasons for adopting such a scheme. First of all, a pragmatic concern for the effectiveness of the system: decomposition into slices facilitates recognition of the particular spatial arrangements which concern us, alignment, transposition and alternation, without overly sensitivity to orientation, as in descriptions in terms of axes.

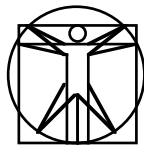
This form of decomposition also allows for the partial investigation of the additive / subtractive procedures through which a location is formed in architectural design. In particular, further subdivision of the slices and/or combination of both horizontal and vertical slicing allows the identification of constraints of such procedures: which sides are the more rigid, which sides extrusion is allowed on, which parts of neighbouring locations might have resulted from the same or similar additive / subtractive operations, and similar indications of formative design operations. Such analysis is more of a complementary investigation than an alternative to current attempts in computer vision to decompose the image of an object into parts that correspond to the natural or canonical parts of the object [Brooks 1981;

Marr 1986; Pentland 1986] as it is an artificial decomposition aimed at recovering aspects of a shape that have to do with the arrangement of its parts.

A basic reason for adopting an artificial decomposition technique is that the Gestalt premise that the whole is not just the sum of its parts appears to hold rather well in architectural plans. On the level of locations, the combination of space continuity and domination of the whole over its details results in fuzzy and ambiguous natural parts. For instance, the precise geometric definition what one might describe as 'the area in front of the window' in a specific room is normally given not so much with reference to the geometry and spatial articulation of the room (with the possible exception of some particular cases, such as bay windows) than to contextual elements, such as room furniture or light coming from outside. An artificial though soundly founded in pragmatic considerations decomposition may avoid such ambiguities while serving its purpose.

# CHAPTER 3

## THE SPATIAL ARRANGEMENT OF LOCATIONS IN ARCHITECTURAL PLANS



### 3.1 Introduction

#### 3.1.1 From recognition of locations to recognition of relationships between locations

Chapter 2 presented an approach to the recognition of locations, the atomic parts of an architectural plan in chap. It discussed the basic feature detection, decomposition and grouping techniques required for identifying locations and outlined the essential framework of precedent approaches and techniques. The present chapter moves from the recognition of atomic parts to the recognition of relationships between atomic parts: given the description of an architectural plan as a set of locations of specific position in the plan and of specific shape, the process described in this chapter attempts to recognize the spatial articulation of the plan, that is, coherent configurations of groups of locations.

Grouping relationships between locations are essentially the same as those between components of a location. The existence of the same formal structures within groups of locations and individual locations of complex shape is hardly surprising since both represent coherent spatial configurations. The correspondence between the grouping of components of a location (cfr. section 2.5) and the grouping of locations in an architectural plan (especially

the first level of grouping — subsection 3.4.1) means that the discussion of precedent and alternative approaches and techniques in Chapter 2 (subsections 2.2.2, 2.4.2 and 2.5.2) and the evolution of the techniques of CHQP from this framework (subsections 2.2.3 and 2.5.3) also applies to the techniques and descriptive structures proposed in this chapter.

Similarities between the processes described in Chapters 2 and 3 are accentuated by the approach which underlies grouping of locations in CHQP. In the absence of typologic studies of classical architecture which could have provided a repertory of location arrangement types, it was considered preferable to adopt a bottom-up (data driven) approach, similar to the one that underlies the majority of computer vision systems, instead of a top-down (model driven) approach (cfr. subsection 3.1.3 and Figure 3-2). The choice of a bottom-up approach means that grouping of locations in CHQP is based on the recognition of specific spatial relationships between locations and also that, as in Chapter 2, domain knowledge (i.e., knowledge of the classical canon) enters group recognition only implicitly, in the form of a set of acceptable grouping relationships and their order of preference.

### 3.1.2 An outline of the chapter

Section 3.1 begins with the correspondences between the discussions of Chapters 2 and 3 and concludes with a review of relevant descriptive formalisms specifically developed for architectural plans, such as shape grammars and rectangular arrangements. This review complements the presentation in Chapter 2 (subsection 2.5.2) of general purpose descriptive techniques from computer vision and is in turn complemented by the review in Chapter 4 (section 4.2) of concepts related to the formal structure and evaluation of the overall spatial articulation of a classical architectural plan.

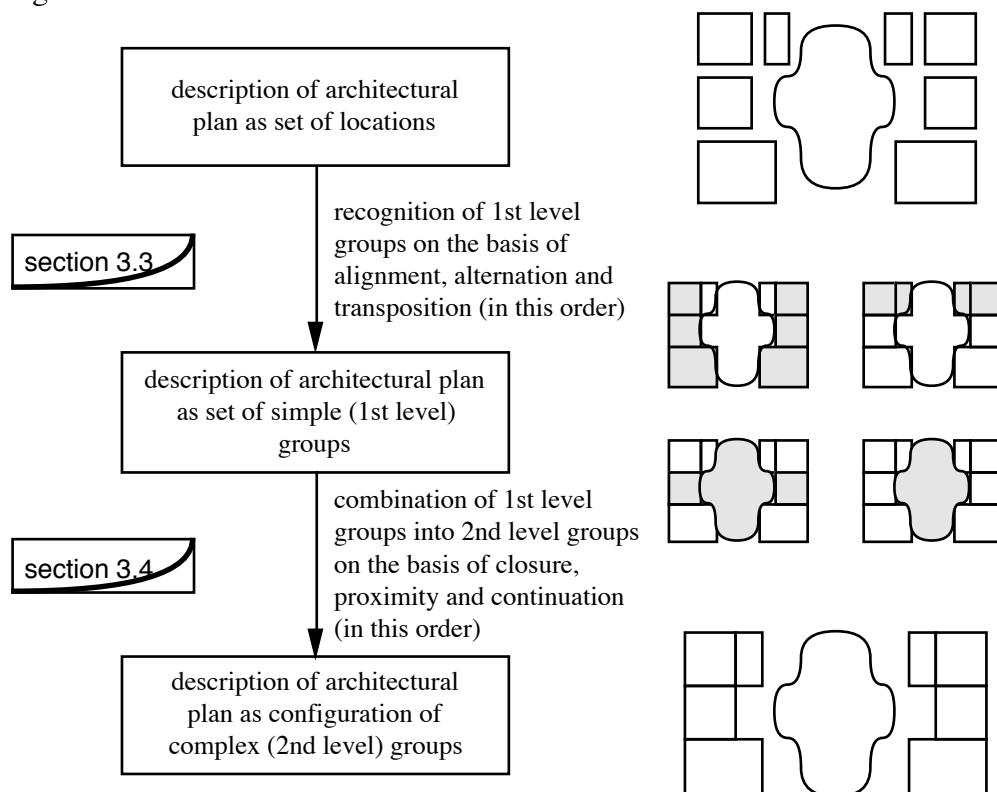
Section 3.2 deals with a major issue of the recognition and representation of grouping relationships, whether they should be considered as bilateral or multilateral. The bilateral case is considered in the form of an adaptation of chain coding to the description of the spatial articulation of architectural plans. The adapted chain coding also serves as an illustration of the local conditions used to detect grouping relationships in both the bilateral and multilateral modes.

Section 3.3 describes the exact types of grouping relationships that concern CHQP and their representation, as well as their relations to Gestalt theory and to representations based on medial axes. Recognition of these relationships in

architectural plans is described in section 3.4. Recognition consists of two levels, recognition of simple (first level) grouping relationships and recognition of complex (second level) grouping relationships. The first level concerns the elementary group types also considered in location shape recognition (cfr. subsection 2.5.4) while the second level combines these into more complex groups which correspond to conventional architectural types of spatial articulation, such as single and double loaded corridor arrangements.

The grouping process described in this chapter results into a description of the plan as a relational network of location groups. In a conventional computer vision system this description would be treated as a more or less definite description of the particular image and would be matched to stored models (i.e., general relational structures) so as to recognize known entities. *chap* follows a modified approach which substitutes the model database with additional levels of analysis and verification. These concern more the structure of the description than its parts and their derivation and are considered in Chapter 4 independently of the grouping process. In anticipation of the discussion and conclusions of the following chapters, the present chapter is rounded off with some interim conclusions (section 3.5) that summarize the particular problems and techniques of the recognition of location groups.

Figure 3-1 presents a schematic outline of the process described in this chapter in relation to the outline of *chap* in Figure 1-1.



**Figure 3-2** A schematic outline of the module of *chap* described in Chapter 3 with respect to the general outline of Figure 1-1

### 3.1.3 Approaches to the representation of architectural plans: the structure of descriptions

Representation of architectural plans in terms of their locations has been rather extensively investigated in CAAD. One critical aspect of such representations which determines to a large degree the structure of resulting descriptions is the nature and kinds of relationships between locations. With respect to that we may distinguish between two basic approaches. The first comprises techniques where these relationships, except for the absolutely necessary one, distance, are not of primary concern and hence do not enter the descriptions of architectural plans. Examples of this category are space allocation and space planning systems [Eastman 1975]. These, while primarily concerned with circulation between locations, may ignore location arrangements which represent very concrete circulation patterns or even circulation spaces themselves [Shaviv 1986, 1987]. The dominant paradigm in these techniques is allocation of locations not primarily in relation to each other but on the basis of a grid, the relationships between grid cells and the standardized components that occupy the cells, as in the system Harness [Hoskins & Bott 1972; Meager 1972, 1973].

The second approach is characterized by the consistent use of a limited number of relationships between locations (or rather interior spaces) which normally result from specific generative operations, as in rectangular arrangements [Stedman 1976, 1983] and shape grammars [Stiny 1975, 1976, 1980; Gips 1975; Stiny & Mitchell 1978a, b]. Such techniques are rather successful in providing a formal structure to the description of architectural plans. This structure is related to a possible *formative history* (to borrow a term from Pentland [1986]) of the plan, that is, a sequence of generative operations that create the particular layout from a limited number of atomic parts, starting with a primeval state, such as a single point in space and a coordinate system with respect to it or a specific area to accommodate the plan.

The generative and classification aspects of the latter approach are of obvious interest but also restrict use of its techniques to particular models and methods of design or to specific plan types. Their major problem is that they relate to a single formative history based on just a few generative operations which cannot accommodate and integrate the various constraints involved in architectural design and the

numerous and largely variable actual formative histories for the same plan type or even the same plan. The rules of a particular shape grammar, for instance, offer the means to encode recursively an architectural plan and hence a reduction of dimensionality but no explicit representation of the procedures and the constraints through which the plan was actually formed. Even if we accept that shape grammar transformation rules are in some sense aggregations of design procedures, it is inevitable that to a large extent the production of a plan layout in this manner is a result of superficial or incomplete domain knowledge, that is in our case, classical architecture. This is evident in the quite restricted application domain for each implementation of shape grammars. So far there has been no evidence for a general shape grammar that could account for a wide spectrum of architectural works, such as a grammar that could produce all classical architectural plans and not just the plans of Palladio's villas. Some may argue that there is not yet a sufficient number of implementations so as to provide the background for unitary investigations but we know from similar attempts in syntactic pattern recognition that "the search for grammars describing large classes of natural patterns has not yet been very successful" [Nevatia 1982, p.22].

What we miss mostly in techniques such as shape grammars and rectangular arrangements is (a) explicit and direct reference to the wider context of an architectural plan, that is, reference not only to patterns observed in Palladian villas but also to the general constraints of the classical canon and even to related human perceptual and cognitive mechanisms, (b) hierarchies of local and global relationships and formative operations which coordinate and control the description and/or production of an architectural plan in a coherent manner, and (c) relative independence of the resulting patterns from corresponding generative operations, that is, the ability to account for different formative histories of the same architectural plan.

In an attempt to investigate the above problems chap adopts an *analytic* rather than a generative approach to the representation of the spatial articulation of architectural plans. The proposed representation does not attempt to generate all possible plan layouts on the basis of an inductively formed rule base but analyses architectural plans on a succession of resolution and specificity levels so as to identify (a) the formal expression of constraints that relate to each level and (b) the precise formal coordinating structures in descriptions of architectural plans, with respect to specific domain theories (i.e., in the initial version of chap, formalizations of the classical canon) and with reference to related human perceptual mechanisms. As a result of this analysis the description of an architectural plan in chap can

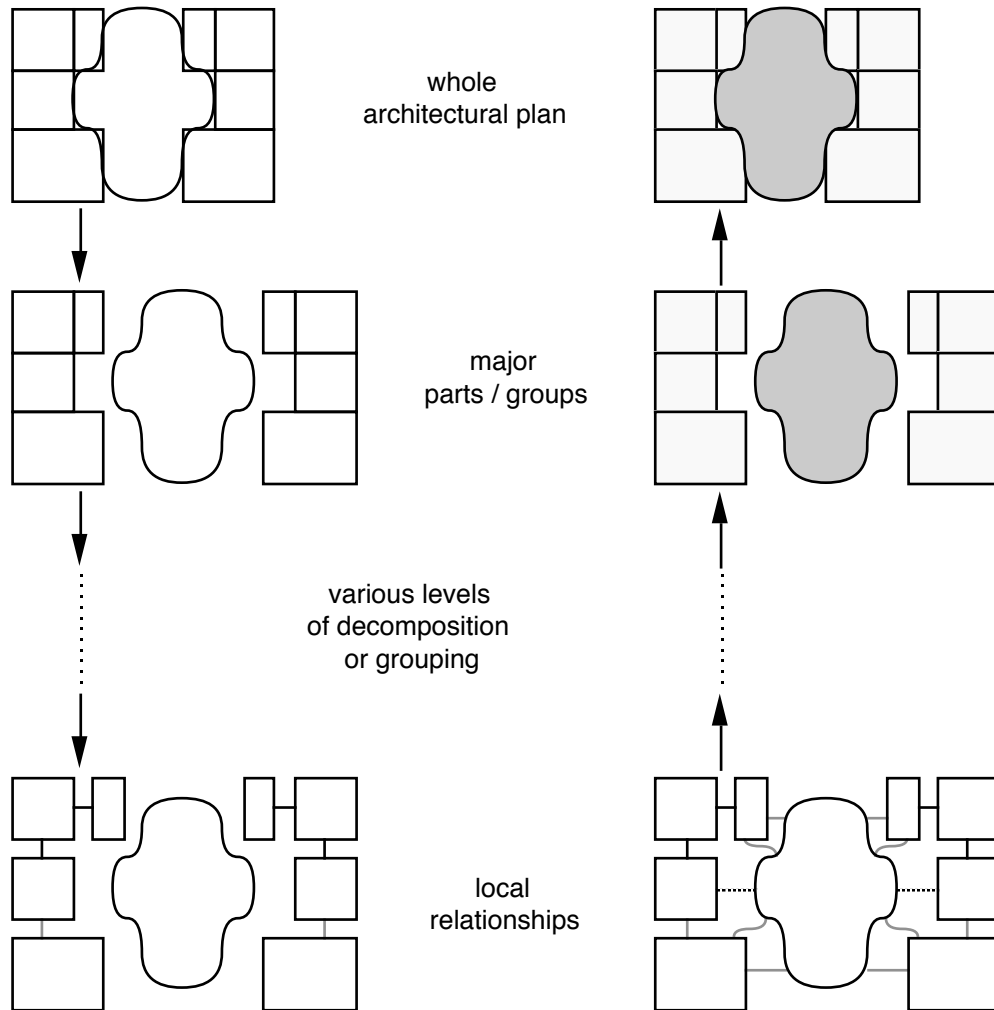
account for a variety of possible formative histories and indirectly correlate these formative histories with the rules and the structure of the classical canon.

A critical problem for the analysis of an architectural plan with respect to general domain constraints, such as the classical canon, is the way the plan is analysed. Given the locations of an architectural plan, i.e., the atomic parts that specify the lowest level of description (the level of highest specificity), there are two basic approaches to the analysis of the geometry of the architectural plan. The one is to start from the whole architectural plan and progressively *decompose* it into parts of increasing specificity on the basis of some formal scheme until the level of atomic parts is reached (Figure 3-2a). The other is to start with the investigation of local relationships between locations and progressively *group* them according to a number of rules into groups of increasing complexity up to the level of the whole plan (Figure 3-2b).

The two approaches are equivalent on a certain abstraction level, as the decomposition of the former can also be expressed as grouping rules of the other and vice versa. Operationally they are not as similar. The decomposition approach is more prone to deterministic recognition of specific formal entities and may impose what appears to be such entities as primary elements of the description even though they are comparatively weak or even 'invent' instances of these entities where they are nonexistent. The grouping approach, on the other hand, tends to present more ambiguities and a greater number of alternatives on each description level. These operational differences suggest that the decomposition approach may result into comparatively less effective techniques and the grouping approach into less efficient ones. In the context of CHQP, that is, in the context of developing an approach rather than a concrete implementation, the grouping approach is therefore preferable.

Another argument for the adoption of the grouping approach is its correspondence to the currently dominant approach to computer vision originally formalized by Marr [1982] and Barrow & Tenenbaum [1978]. In this, vision proceeds from the level of local image features (atomic primitives) to a succession of higher levels which attempt to recover known entities and their parts in configurations of features. Each level is defined on the basis of preceding levels in a bottom-up (data driven) manner. The strong parallels between this process and the grouping approach allow the straightforward adaptation of a number of computer vision techniques to the recognition of the spatial articulation of architectural plans. The availability of existing techniques, the consistency and coherence of the methodologic framework derive from and the affinities of

that framework with the approach of CHAP are primary reasons for adopting the grouping approach at this stage of CHAP because they compensate for the lack of typologic studies of classical architecture which could have provided the prerequisites of the decomposition approach, a repertory of location group types and their relationships with the whole architectural plan.



**Figure 3-2** *A schematic comparison of the decomposition and the grouping approaches to the analysis of architectural plans*  
*a (left): schematic outline of the analysis of an architectural plan by decomposition;*  
*b (right): schematic outline of the analysis of an architectural plan by grouping*

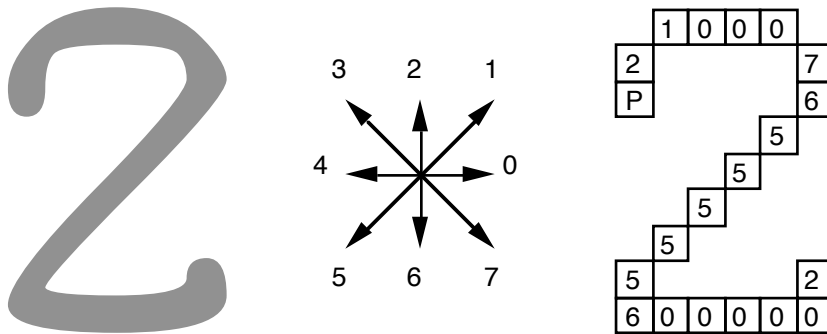
## 3.2 The nature of grouping relationships in architectural plans

A critical aspect of a bottom-up recognition system is how the atomic primitives are linked together to form higher-level entities. The presentation of the linking structure can be subdivided into two: the *types of relationships* between primitives, described in the following section (3.3), and the *nature* of these relationships, which is the topic of this section. This distinction is helpful for understanding the various grouping operations in *CHAP*, the representation of resulting groups and connections with human perception. It also allows to differentiate between the local conditions used for the recognition of a grouping relationship and the overall representation of the relationship.

### 3.2.1 An example of bilateral relationships: an adaptation of chain coding

Scene analysis and Artificial Intelligence have traditionally concentrated on descriptions based on bilateral relations between descriptive primitives and accordingly resulting combinations. A simple though characteristic example of such representations is *chain coding* [Freeman 1961]. The present subsection introduces an adaptation of chain coding to the description of the spatial articulation of architectural plans that carries forward all conceptual and operational characteristics of conventional chain coding. This adaptation is capable of representing the relationships of alternation, transposition and alternation between rectangular components of a location we have considered in the Chapter 2, as well as similar relationships between locations.

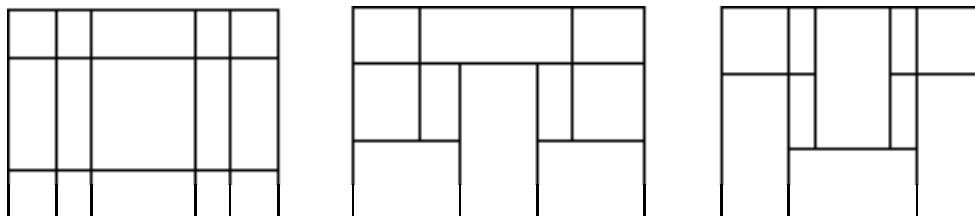
Chain coding is a basic and efficient manner of describing digital lines in terms of topologic relationships between their elements. It relies on the assignment of a standard label to each of the neighbours of a pixel and operates by recording the coordinates of a starting point (normally one of the endpoints) and then the label which corresponds to the relative position of each successive pixel (Figure 3-3). The resulting description is a string that is more easily manipulated and requires less storage space than the original two-dimensional image array or the ordered lists of coordinates (absolute or relative) of its pixels.



**Figure 3-3** Chain coding

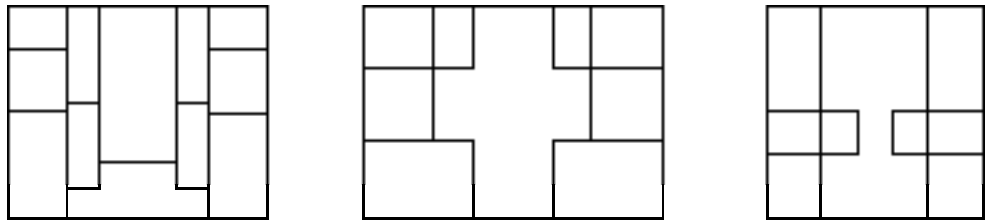
left: a two-dimensional curve; middle: chain code for an eight-connected rectangular grid; right: the curve decomposed into accordingly labelled discrete parts — the resulting string which describes the curve is P2100076555556000002 (P is the starting point)

Chain coding is applicable in principle to linear pixel sequences with few if any branching points. Adaptation of chain coding to the description of contiguity relationships between locations of an architectural plan involves accounting for (a) elements of variable shape and size because not all locations of an architectural plan have the same shape and size, unlike the pixels of a digital picture, and for (b) multiple groupings of each element because, while a line is a sequence of pixels where each pixel may belong to just one group of pixels (i.e., a single line segment — unless it is a branching point), an architectural plan is an array of locations and hence each location may fit into more than one group.

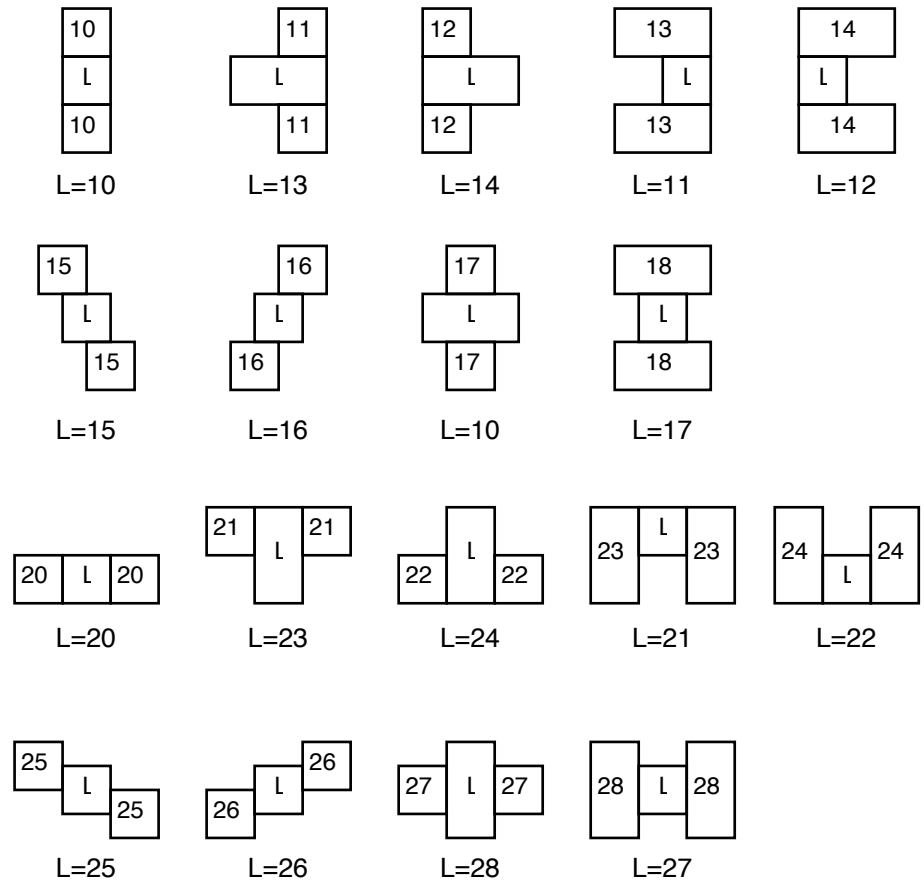


**Figure 3-4** Totally orthogonal architectural plans with rectangular locations

These problems are scaled down by the totally orthogonal environment of CHAP and by considering initially only architectural plans whose locations have all rectangular shape, as in Figure 3-4. This environment represents a lowest common denominator for a variety of floor plan types and architectural styles and should be considered as a stepping stone to covering more complex geometries. In fact, later in this subsection we shall extend to totally orthogonal architectural plans which contain locations of complex shapes, as in Figure 3-5.



**Figure 3-5** *Totally orthogonal architectural plans with locations of complex shapes*



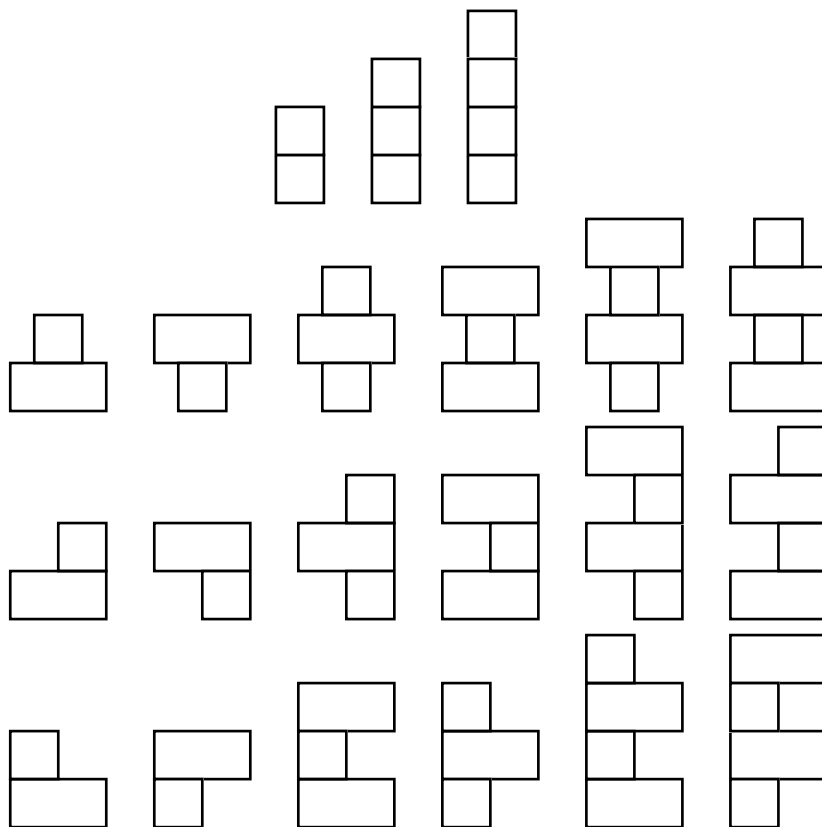
**Figure 3-6** *An adaptation of chain coding to the description of contiguity relations between rectangular locations*

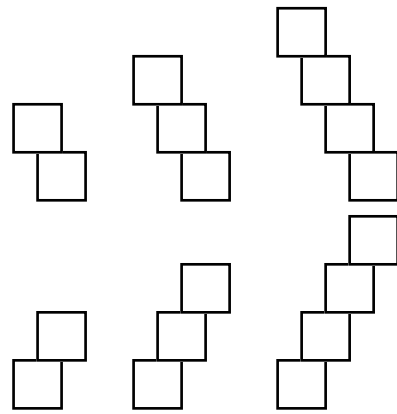
Figure 3-6 shows all possible size and position classes of rectangular locations that can be contiguous to a rectangular location. The labels attached to these classes are subdivided into two major groups. Labels 10 to 18 denote a vertical direction of articulation while labels 20 to 28 denote a horizontal direction of articulation in the resulting group of locations. Unlike conventional chain coding, the one label applies to two neighbours which have qualitatively the same relationship with the reference location as they occupy symmetrical positions relative to it. Another deviation is that labelling is reciprocal as both locations in a contiguous pair are labelled. This means that (a) there is no starting point in the resulting description and that (b) each location is initially marked with as many labels as the number of its contiguous locations. This necessitates the establishment of a preference

order for selecting the proper label.

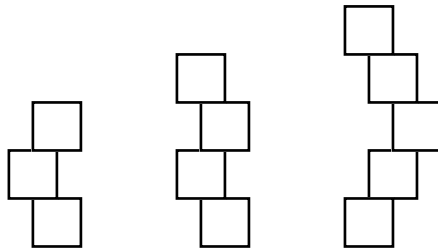
The preference order relies on the group type suggested by each label. Figure 3-7 shows examples of vertically articulated group types. (Horizontally articulated group types are essentially similar and therefore are characterized by an identical preference order.) The group types are characterized a different preference order in each architectural formal system. In general, a group of type 10 (20 in the horizontal direction) is the highest in preference order; then come either type 17/18 (27/28) or types 11/13 and 12/14 (21/23 and 22/24), depending upon whether alignment is more significant than axial symmetry. Usually types 15 and 16 (25 and 26) come last. A preliminary investigation of studies on classical architecture, including Stiny & Mitchell [1978a, b] and Freedman [1987], suggest that can an acceptable preference order in classical architectural plans is 10, 17/18, 11/13 and 12/14, 15 and 16 (vertically); 20, 27/28, 21/23 and 22/24, 25 and 26 (horizontally).

The collection of group types in Figure 3-7 is not exhaustive. One might suggest more group types, such as those in Figure 3-8. In this initial version of adapted chain coding it was considered preferable to treat such groups as combinations (aggregations or fusions) of the generic types of Figure 3-7, so as to restrict the collection of primitive group types to the absolutely essential ones and thus keep the size of the corresponding rule base as compact as possible.





**Figure 3-7** *Examples of vertically articulated group types with up to four locations*  
*from top to bottom: group type 10; group type 17/18; group type 11/13;*  
*group type 12/14; group type 15 (left), group type 16 (right)*



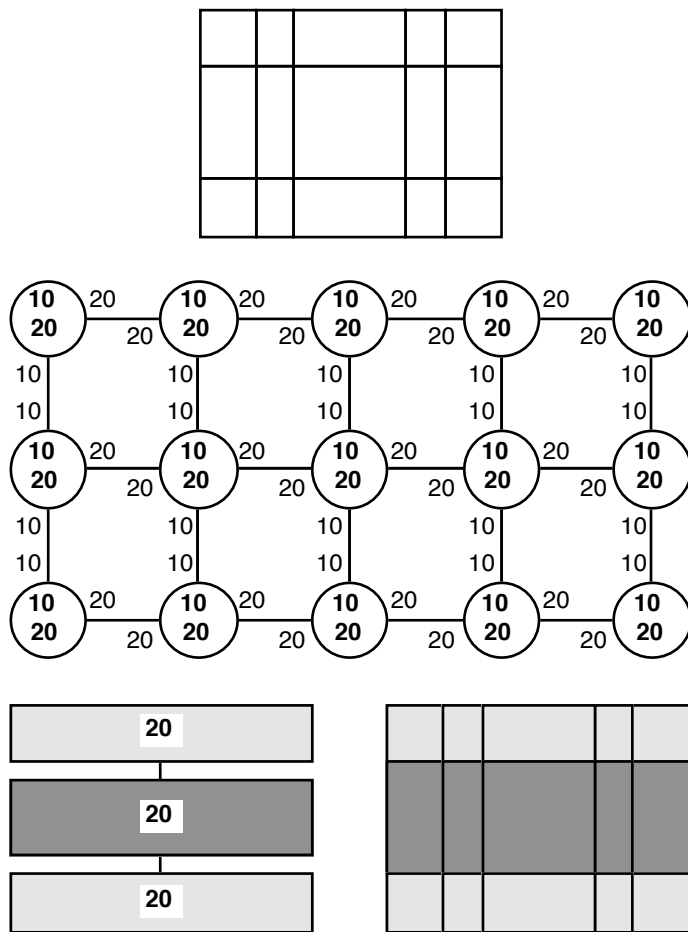
**Figure 3-8** *Additional group types which can be considered as aggregations or fusions of the group types in Figure 3-7*

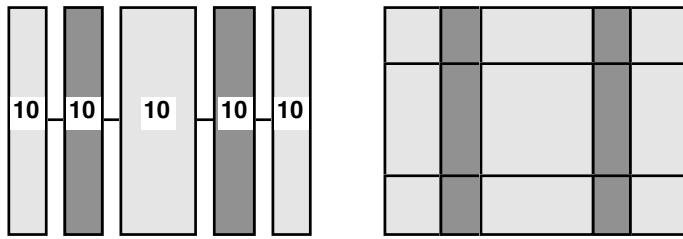
The preference order for groups determines the preference order for location labels. Label 10 is generally preferred to all others which denote vertical articulation. The order of the rest depends upon the same domain constraints as the order of group types. In classical architectural plans the preference order of location labels is 10; 17, 18; 11, 12, 13, 14; 15, 16 for vertically oriented groups and 20; 27, 28; 21, 22, 23, 24; 25, 26 for horizontally oriented groups (labels separated by commas are equivalent). Labels which denote horizontal articulation are considered to be incommensurate to labels which denote vertical articulation. Therefore, of the labels attached to each location two are retained in the end: one denoting the preferred horizontal and the other the preferred vertical group type. The resulting description of each architectural plan is dual in that it comprises a configuration of horizontally articulated groups and a configuration of vertically articulated groups.

Preference for a location label also depends on grouping constraints which essentially amount to that in the description of an architectural plan there should be as few possible solitary locations, that is, locations that form a group by themselves. For example, if a location is labelled with two completely equivalent codes which denote different group types, such as 11 and 12, preference of the one over the other relies on whether it results into solitary locations in

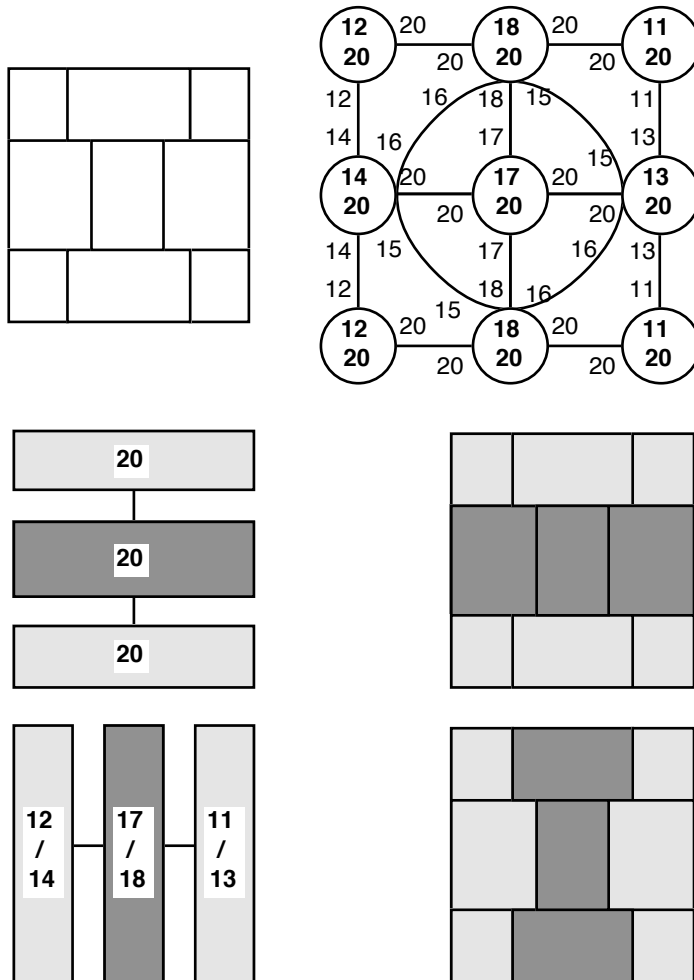
the final configuration of groups. In a sense, grouping constraints apply more to the level of alternative configurations of groups than to the level of alternative labels of individual locations and therefore constitute a form of feedback mechanism which confirms or overrides the initial selection of a location label.

Figure 3-9 depicts the application of adapted chain coding to the plan of Palladio's Villa Thiene. Contiguity is recorded in a graph where nodes represent locations and links contiguity relations between locations. A link is marked with the adapted chain coding labels which correspond to the particular relationship, while a node is marked with the preferred horizontal and vertical labels for the particular location. On the basis of the preferred location labels two group configurations are derived, one of horizontally articulated and one of vertically articulated groups. In the corresponding diagrams blocks represent location groups and links contiguity between groups. In Villa Thiene three horizontal groups of type 20 and five vertical groups of type 10 are identified.





**Figure 3-9** Application of adapted chain coding analysis to the architectural plan of Palladio's Villa Thiene  
 from top to bottom: schematic floor plan; contiguity graph; horizontal and vertical grouping



**Figure 3-10** Application of adapted chain coding analysis to the architectural plan of Palladio's Villa Ragona

Villa Thiene is an elementary case for the application of adapted chain coding as it involves only two kinds of contiguity relationships. Conflicting labels occur at no node and the preference order need not be employed. Still, Villa Thiene is significant for the discussion of adapted chain coding because its architectural plan corresponds fully to the “geometrical pattern of Palladio’s villas” [Wittkower 1952, p.65], that is, it can be considered as the design where Palladio’s rules of geometric arrangement have been applied more extensively.

To anticipate the discussion of the classical canon in Chapter 4 (subsection 4.2.2), this 5 x 3 grid pattern is

considered as an essential aspect of Palladian composition. Wittkower [1952, Part III, Figure 8] describes the plans of Palladio's villas with reference to this pattern. In the Palladian grammar [Stiny & Mitchell 1978a, b] the first stage of the composition of an architectural plan, grid definition, results invariably into a 5 x 3 or a 3 x 3 grid pattern. Ackerman [1977, p.60] suggests that there is a uniform triadic schema which underlies plans and elevations, a triadic coordinating structure which supports and facilitates the extensive and unified use of proportionalities and also conforms to the traditional centralized arrangement of the typical Venetian house.

The concept of a coordinating structure is further investigated by Tzonis & Lefaivre [1986] for the whole of classical architecture. According to Tzonis & Lefaivre subdivision of a classical building is constrained by the *tripartition schema* which seems to be embodied in the generic nine squares pattern,

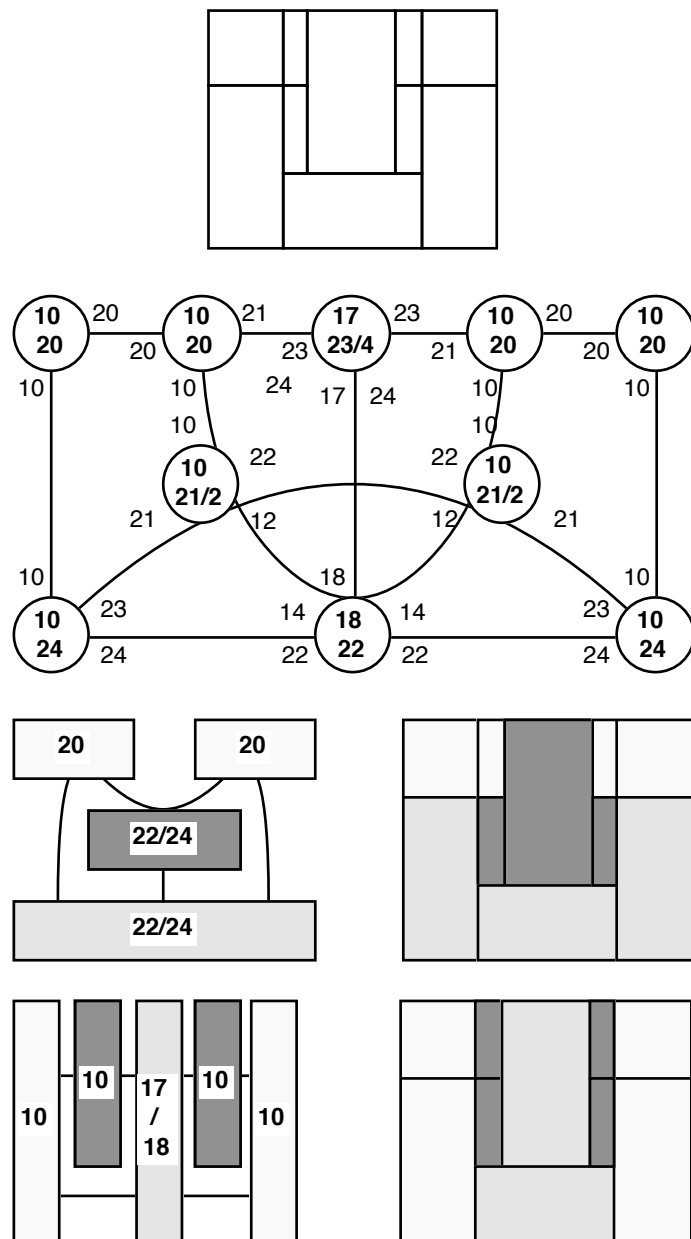
a	b	a
b	c	b
a	b	a

i.e., the 3 x 3 grid pattern, from which abstract plan layouts are derived in a top-down manner through deletion, fusion, addition (or repetition) and embedding of parts [Tzonis & Lefaivre 1986, p.24]. For example, the transformation of the 3 x 3 into the 5 x 3 pattern is performed by the addition of two identical columns:

a	b	a			a	d	b		
			d	a					
b	c	b	(ADDITION:)		b	e	c		
			e	b					
	a	b	a		a	d	b	d	a

Although there is more in tripartition and the classical canon (cfr. subsections 4.2.2, 4.4.3 and 4.4.4), it is sufficient for the moment that tripartition, in the form of the 3 x 3 or the 5 x 3 grid pattern, is a fundamental aspect of a classical architectural plan. The configuration of location groups recognized by adapted chain coding should conform to either of the two patterns, which in the case of Villa Thiene (Figure 3-9) is beyond any doubt.

Slightly more complicated is the case of Villa Ragona (Figure 3-10) where many locations are labelled with three alternative vertical labels. Selection of the proper label on the basis of the preference order discussed previously results into a group configuration that satisfies tripartition constraints.

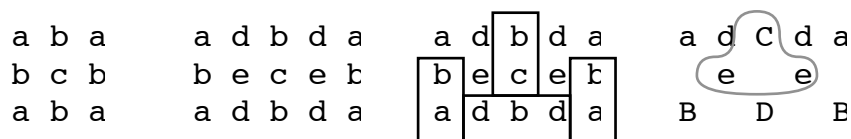


**Figure 3-11** Application of adapted chain coding analysis to the architectural plan of Palladio's Villa Poiana labels such as 21/2 in the contiguity graph denote that the location has both labels 21 and 22 attached to it as initially there can be no preference for the one over the other

More extensive application of the preference criteria is required for Villa Poiana (Figure 3-11) which contains also locations with alternative horizontal labels. The resulting vertical grouping confirms our intuitive expectations but the horizontal grouping appears to be counterintuitive with respect to the middle group (type 22/24). This group consists of three locations which appear to hold stronger relationships with locations of other groups than with each other.

Counterintuitive groupings may result in adapted chain coding (another example being the vertical grouping of Villa Zeno in Figure 3-13) because the instances of the elementary

group types of Figure 3-7 do not combine into more complex groups. This means that adapted chain coding is ultimately an initial level in a sequence of groupings and formal interpretations equipped with feedback and recursion mechanisms. Still, the configuration of location groups returned by adapted chain coding should not be considered as potentially erroneous nor as an incomplete preliminary structure that should be tolerated and not rigorously analysed. On the contrary, it should be treated as a serious *analytic* device that facilitates recognition of underlying implicit relationships which may not be obvious in intuitive examinations of an architectural plan.



**Figure 3-12** *The articulation of locations in Villa Poiana explained in the manner of Tzonis & Lefaivre [1986]*

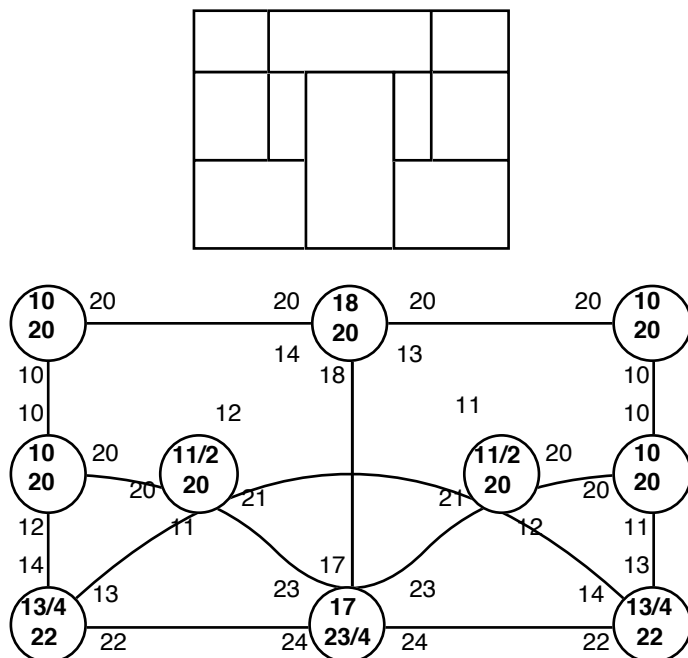
*from left to right: the 3 x 3 grid pattern; the 5 x 3 grid pattern derived by the addition of two columns of locations; fusion of grid cells to produce the pattern of Villa Poiana; the resulting pattern with the group under investigation (see text) circumscribed by a dotted line*

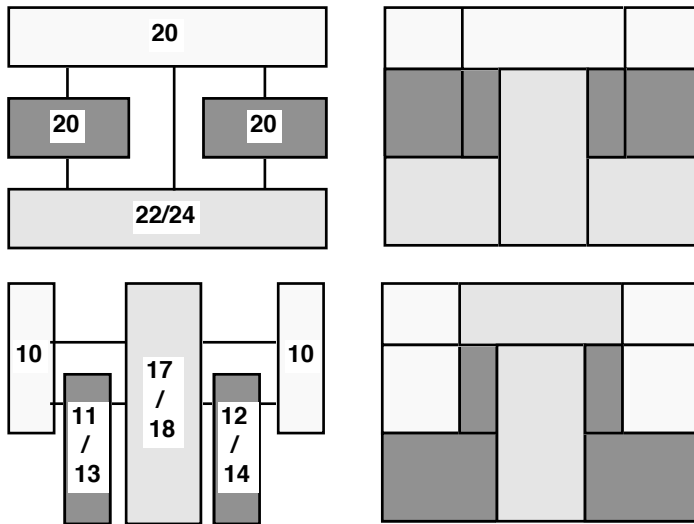
Investigation of the grouping of Villa Poiana (Figure 3-11) as an expression of underlying rules and relationships suggests that it is consistent with the constraints of classical canon. Derivation of the spatial articulation of Villa Poiana in the manner of Tzonis & Lefaivre (Figure 3-12) suggests that there are good reasons for grouping locations *e*, *C* and *e* (to use the symbolisms of Figure 3-12) together — at least as good as for grouping *C* with the locations of the top row (i.e., with *a*, *d* and *d*, *a*): *C* is the product of the fusion of *b* and *c*, two cells that belong respectively to the top and the middle row. Since location *C* holds equivalent alignment relationships with the locations of the top and of the middle row, both *adCda* and *eCe* are in principle equally acceptable groupings. Within the context of adapted chain coding *adCda* is, firstly, a complex group (a combination of types 20 and 21/23) and, secondly, its preference over *eCe* would leave the two *e* locations solitary despite their obvious affinities. Preference of *eCe*, does not leave any solitary locations and is subsequently accepted. Similar considerations underlie the justification of the two vertical groups *eA* over the two *abA* groups in the architectural plan of Villa Zeno (Figures 3-13 and 3-14).

The horizontal grouping of Villa Poiana (Figure 3-11) and of Villa Zeno (Figure 3-13) shows that each of the three (or five) parts in the tripartition schema may comprise more than one groups of locations. The top row of Villa Poiana and the middle row of Villa Zeno comprise two distinct groups of locations. Intuitively each of these pairs is treated

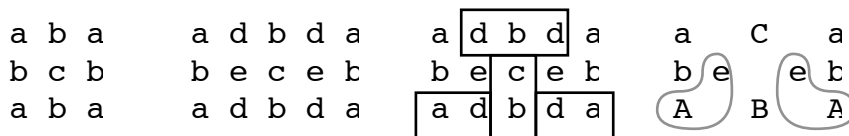
as the collection of the fragments of a single group that has been interrupted by a lacuna, that is, in the system of Tzonis & Lefavre, the fusion of one or a couple of the grid cells of that row with cells of other rows. In Palladian villas such pairs are symmetrical with respect to a central axis and therefore can be correlated merely on the basis of their position relative to the middle group of the plan. A more general formulation of this correlation would involve identification of *continuation* (in the sense of Gestalt theory —cfr. subsection 3.3.1), for instance in the form of common or connected alignment fronts (cfr. subsection 3.4.2).

In Palladian villa plans a location may be rectangular or cross, T- or double-T- shaped. Only one location in each plan may be of complex (i.e., not rectangular) shape. This location must be centrally located (bisected by the symmetry axis of the plan) [Stiny & Mitchell 1978a, p.9]. Augmentation of the labelling scheme of Figure 3-6 with the possible types of contiguity relations between each complex shape and its rectangular neighbours still results into a relatively compact labelling system but only if we restrict the applicability of adapted chain coding to Palladian villas only. Otherwise it is necessary to consider contiguity relationships between all complex shapes and also all possible variations of each shape. Inclusion of all relationships would result into a rather large, cumbersome and inefficient labelling system.





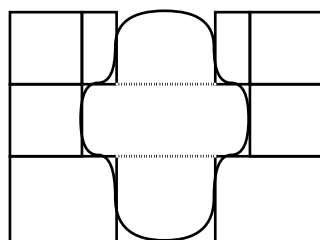
**Figure 3-13** Application of adapted chain coding analysis to the floor plan of Palladio's Villa Zeno

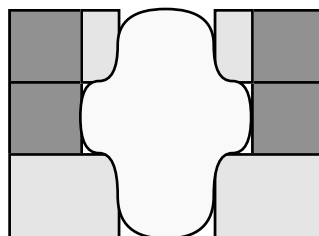
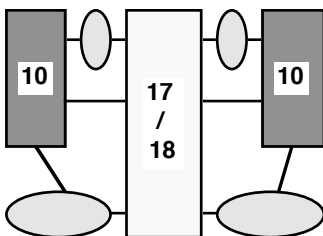
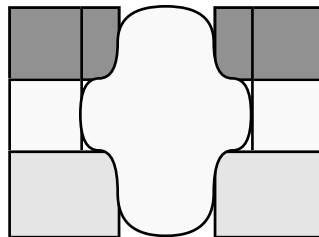
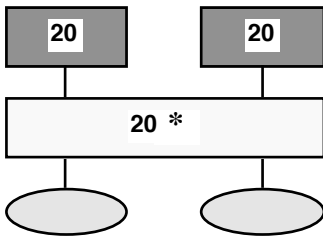
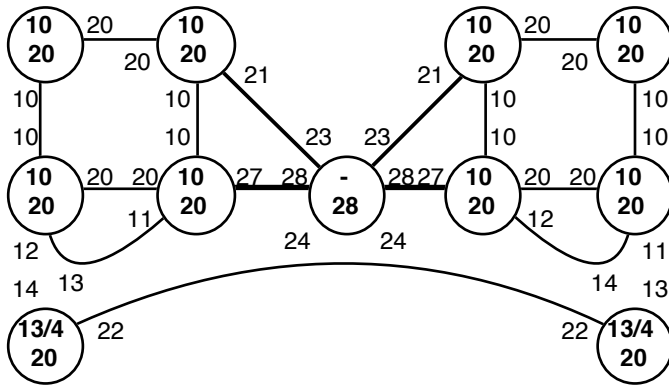
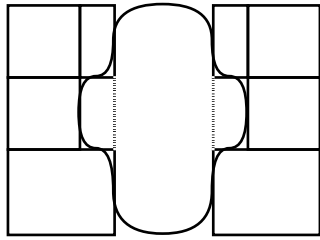
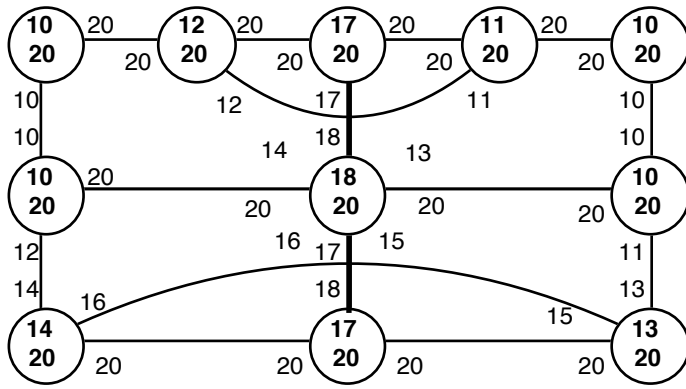


**Figure 3-14** The articulation of locations in Villa Zeno explained in the manner of Tzonis & Lefaivre [1986]

An alternative approach is to decompose locations of complex shapes into rectangular components as in Chapter 2 (section 2.5) and consider contiguity relations between these components and rectangular locations. This allows the use of the labelling system of Figure 3-6 for all rectangular architectural plans, simply by augmenting the preference criteria for grouping. This extension of adapted chain coding is applied to the plan of Villa Foscari (Figure 3-15).

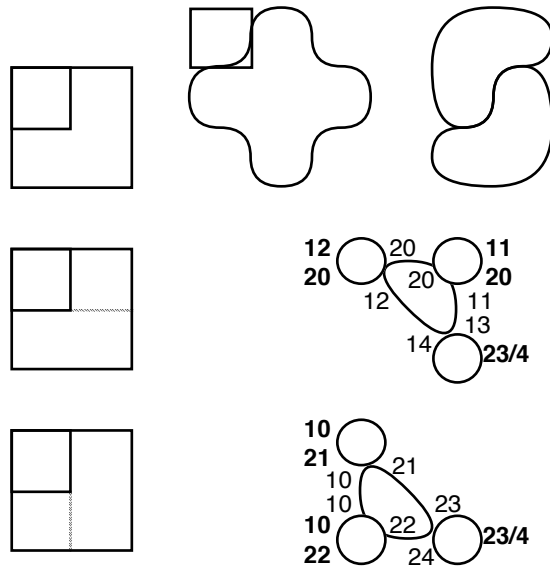
In order to explore the adopted decomposition technique more comprehensively, the central cross-shaped location of Villa Foscari is segmented alternatively into both horizontal and vertical slices (Figure 3-15) because horizontal slices are seemingly better suited to the recognition of vertically articulated groups and vertical slices to the recognition of horizontally articulated groups. However, in the case of Villa Foscari at least, both horizontal and vertical slices result into the same configuration of groups of locations.





**Figure 3-15** Application of adapted chain coding analysis to the floor plan of Palladio's Villa Foscari  
 opposite (from top to bottom): schematic floor plan under horizontal slicing; contiguity graph under horizontal slicing (fat links represent contiguity between components of the same location); this page: schematic floor plan under vertical slicing; contiguity graph under vertical slicing; horizontal grouping (oval nodes represent solitary locations, an asterisk after the group type indicates that the group contains locations of complex shape); vertical grouping

The preference criteria for grouping are augmented with the addition of the following rules: (a) The components of a location must all belong to the same group. (b) A rectangular location may be grouped with a location of complex shape only if it holds the same relation to all components it is contiguous to, i.e., receives the same label as a result of its contiguity to these components. The purpose of the first rule is to preserve the integrity of locations of complex shape which is obviously stronger than any relation of any of its components to other locations. The purpose of the second rule is to allow minimal deviation from the simple, generic group types used in the present version of adapted chain coding. It excludes relationships of closure (Figure 3-16) but also allows to ignore shape variations among the members of a group, as in the middle horizontal group of Villa Foscari (Figure 3-15), where a cross-shaped location is grouped with two rectangular locations to form a group of type 20.

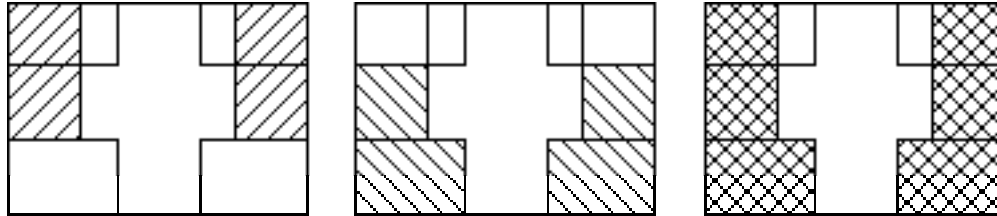


**Figure 3-16** Closure

*top row: examples; middle row: analysis of closure in terms of contiguity relations on the level of component rectangles (horizontal slicing); bottom row: analysis of a case of closure in terms of contiguity relations on the level of component rectangles (vertical slicing)*

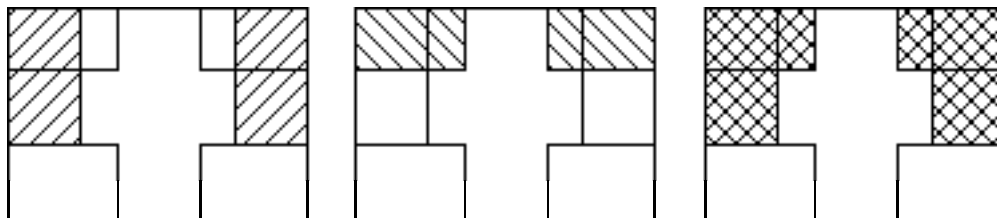
Although recognition of tripartition is achieved in the application of adapted chain coding to Villa Foscari, the resulting group configurations are not as compact as in the previous cases, especially in the configuration of vertical groups where most locations remain solitary. This is not due to the preference for decomposition over augmentation of the labelling system with complex shapes. Even if relationships such as closure (Figure 3-16) were included in the labelling system we would still fail to come up with the intuitively acceptable configuration of two U-shaped groups of rectangular locations that flank the central cross-shaped location. If, on the other hand, the grouping rules were

relaxed in a manner that would allow for fusion of related groups (such as 10 and 11/13 or 12/14) and also of intersecting horizontal and vertical groups of essentially the same type (such as 10 and 20), then recognition of the two flanking groups would be achieved (Figures 3-17, 3-18).

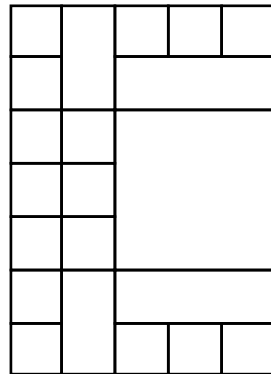


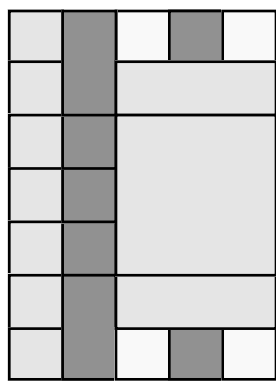
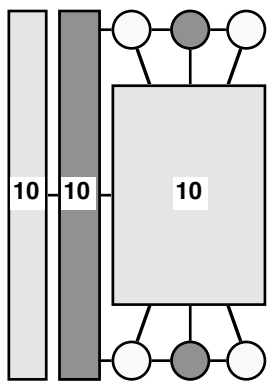
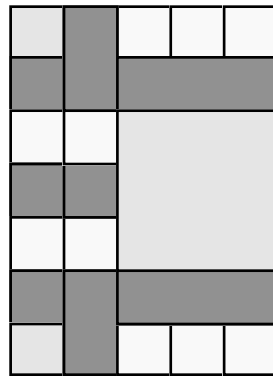
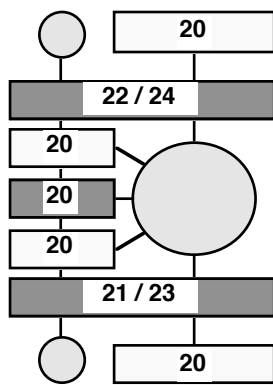
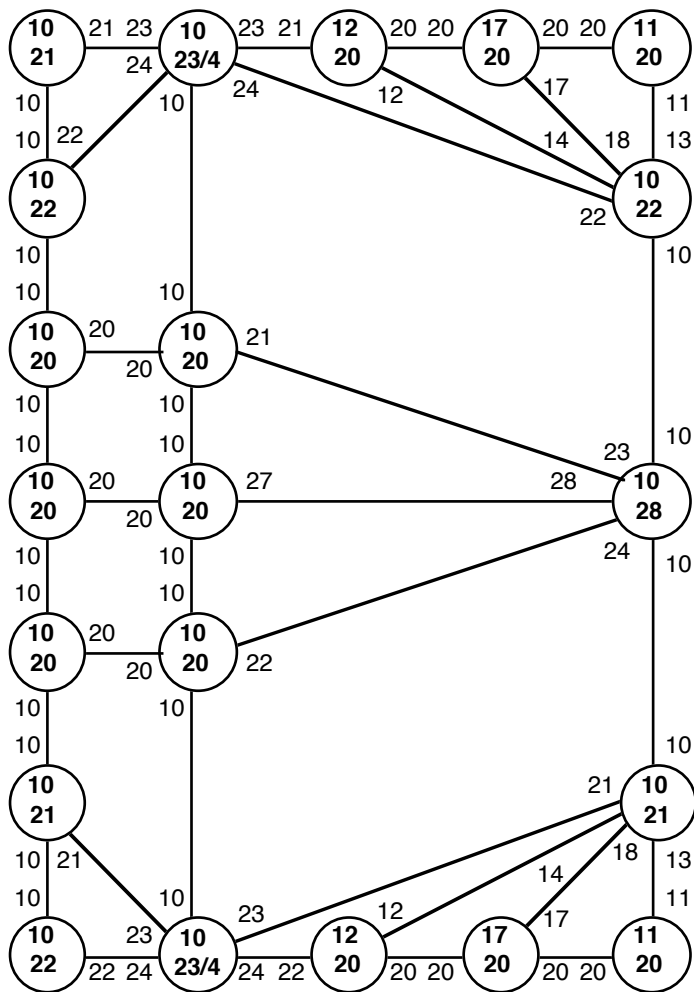
**Figure 3-17** *Combination of groups with the same direction of articulation in the case of Villa Foscari*  
the two groups of type 10 (left) combine with the groups of type 11/13 and 12/14 (middle) to form one group of type 10/11/13 and one of type 10/12/14 (right)

The intuitive acceptance of configurations that may be obtained by such relaxation of the grouping rules indicates that the level of grouping achieved by adapted chain coding is only an initial one that should be followed by additional levels that account for more complex or abstract relationships and group types. Closure is a clear example of this necessity, as a definition of closure in terms of contiguity relationships between rectangular components inevitably involves the combination of relationships of different orientation of articulation (Figure 3-16).



**Figure 3-18** *Combination of groups with different direction of articulation in the case of Villa Foscari*  
the two groups of type 10 (left) combine with the groups of type 20 (middle) to form two groups of type 10/20 (right)





**Figure 3-19** Application of adapted chain coding analysis to a larger architectural plan than Palladian villas

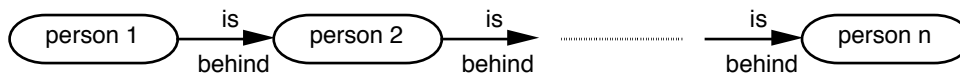
The fundamental weakness of adapted chain coding lies not in the level of grouping it achieves (as further levels can be easily added to it) but in that the essentially *multilateral* relations between locations of a group are determined and described by *bilateral* relations, that is, relations between pairs of contiguous locations. Besides that spatial relations between such pairs are heavily constrained by their immediate context on a multilateral scale, the description of multilateral relationships in terms of bilateral ones is, as we shall see in the following subsection (3.2.2) and in section 3.3, contrary to our perceptual and cognitive mechanisms. Representations which rely on bilateral relationships for the description of multilateral relationships present certain problems when applied within chap. These become more evident in larger or more complex architectural plans than Palladian villas, such as that in Figure 3-19.

### 3.2.2 Multilateral interpretation of binary relationships

Figure 3-19 provides a good illustration of some of the arguments against the use of bilateral relationships for the representation of multilateral ones. The more critical ones arise when we compare the structure of the description which results from the application of adapted chain coding with our perceptual and cognitive abilities and tendencies. For example, once the viewer has been accustomed to adapted chain coding conventions, his intuitive interpretation of the contiguity graph shift from binary to *multilateral*. When he sees that the vertical label of the top left location is 10, a label which denotes vertical alignment on both sides, he scans quickly downwards that column so as to establish the extent of the alignment group rather than check laboriously the sequence of location labels. He may not be able to tell the number of the locations that are aligned or he might make minor mistakes (for instance include a location that is not aligned), but the efficiency and abstraction of this intuitive interpretation cannot be reproduced as long as the relations between locations are only binary in nature.

Similar recognition of multilateral relationships prevails perception in all cases. If we ask viewers to identify the leftmost column of aligned locations (i.e., the leftmost group of type 10) in the abstracted binary image of Figure 3-19, we can expect that they will all perform equally well and fast, although afterwards only the most observant will be able to recall the precise number of locations in the column while most will simply guess a probable number. On the other hand, all will remember that

the group spans the entire length of the plan.



**Figure 3-20** *Description of a queue in terms of bilateral relationships between the people in the queue*

The same applies to the description of entities such as a queue. In most AI and computer vision systems the description of a queue would be in terms of the people in it and bilateral relationships between them (Figure 3-20). Such descriptions are of limited utility because they concentrate on all the members of the queue and partial relationships between them. An intuitive description of the same queue would concentrate on exactly the aspects that are only implicit in Figure 3-20, such as the overall shape and position of the queue. The number of people in it would be immaterial, except for certain thresholds, and members of the queue would be mentioned individually only if they occupied a characteristic position (such as the first or last in the queue) or deviated from the overall structure (for instance someone not aligned in position with the queue shape).

Similarly, a double-loaded corridor in an architectural plan is not described in terms of the spatial relations between each location and its neighbours, but as two rows of locations which flank the corridor location, that is, two multilaterally defined groups and their relations to a solitary location. As we can see from the examples used by Gestaltists (cfr. subsection 3.3.1) the tokens of a linear arrangement can be points, all sorts of lines and shapes with area, and still form a line without losing their integrity and without altering radically the structure and perception of the line.



**Figure 3-21** *The overall shape of a queue  
the stimulus may be perceived as a single entity with overall characteristics such as shape although it consists of distinctively discrete parts*

The description of a queue as a relational graph of all its members is not only counterintuitive but also pointless if we are interested in general characteristics of the queue rather than who is in it and where. The realization that the manipulation of such entities as a whole demanded more abstraction appears to have motivated a fair proportion of AI and computer vision research, as far as one can judge from

the many approaches and techniques developed for abstracting collections of parts into higher-level entities. Most of them are model-based, that is, they rely on templates derived from prior knowledge. These templates are matched to parts of the relational structure which describes the whole of a scene or situation [Brooks 1981; Brown & Ballard 1982, p.8, Fig.1.4d & Part III; Hanson & Riseman 1978; Weymouth et al 1983], in an attempt to abstract groups of lower level entities into known higher level entities. This form of abstraction alleviates the problem to a certain degree as it allows the collective description and manipulation of all members of a coherent group. However, it does not resolve several fundamental issues, as it does not allow, for instance, direct recognition of the shape of a queue.

The potential of multilateral relationships goes beyond linking together all pairs resulting from bilateral relationships. Recognition of the queue as a coherent whole with a particular shape, texture and position in space rather than as a number of distinct people is similar to recognition of the figures Gestaltists made of dots and other abstract shapes (Figure 3-21). From that perspective, the bilateral relations between the people in the queue are just local expressions of the multilateral relationships with respect to local constraints such as proximity thresholds and rotations of orientation reference frames.

In computer vision transition from multilateral to bilateral relationships has been indirectly addressed by the use of multiple resolution levels for the same image, as in Marr [1982]: while on a certain resolution level a human figure, for instance, is perceived as a single, integral cluster of picture elements and described by a single generalized cone, finer resolution levels identify more than one clusters in the same area of the image and hence suggest that the initial generalized cone actually comprises several components which are also described by generalized cones. Similarly, on a coarser resolution level the queue of Figure 3-21 would appear as a continuous curve whose characteristics essentially correspond to the multilateral relationships which characterize the grouping of the entities that comprise the curve on the resolution level of Figure 3-21. Even though I am inclined to accept the criticism of this approach by Witkin [1986] and Pentland [1986], there can be little doubt that the use of multiple resolution levels can be more effective in decomposing higher level entities into their components or in grouping entities of a certain level into higher level ones than approaches that rely on descriptions of a single resolution level, even when they employ sophisticated structuring and matching techniques for grouping and recognition [Minsky 1975; Winston 1975].

Unfortunately, the use of multiple resolution levels as in Marr [1982] or Witkin [1986] for aggregating locations

into groups is not directly applicable to CHAP, despite that it resolves the problem of subordination of bilateral relationships to multilateral ones. The main reason is that this technique would reduce *reading* an architectural plan to *seeing* an architectural plan. As Arnheim [1969, p.13] has argued, “the cognitive operations called thinking are not the privilege of mental processes above and beyond perception but the essential ingredients of perception itself”. If we treat an architectural plan as an abstract visual pattern and not as a representation of a spatial structure, we ignore all the cognitive elements that play a significant role in, for instance, dealing with the variability of location shapes and sizes and also with the multiplicity of possible alternative groupings for each location that may be perceptually (visually) equally acceptable. As we shall see in Chapter 4, our interpretation of architectural plans and other similarly symbolic images ultimately involves very concrete and coherent cognitive constraints which may or may not agree with our perceptions. A clear example of this disagreement are alphabets, where pictorial similarities between different characters do not correspond to phonetic similarities. As a result, classification of characters in the Latin alphabet depends on phonetic constraints that cannot be represented in the visual appearance of the characters. This, however, poses no problem to any literate person.

In view of the similarly cognitive structures which constrain reading architectural plans CHAP reproduces the results of the use of multiple resolution levels through the use of a succession of grouping levels. The first of these levels concerns recognition of multilateral relationships between locations. Each of the subsequent levels is produced by the correlation and connection of the resulting groups. On each grouping level each group is represented by the (multilateral) relationship that binds its locations and the whole plan by the configuration of such relationships. This succession of grouping levels is essentially equivalent to a system of multiple resolution levels. The only difference lies in that in the case of multiple resolution levels the task is to correlate a recognized arrangement with arrangements of other levels on the basis of their position in the image, while in the case of multiple grouping levels the task is to identify the form of an arrangement of a particular level through the correlation of arrangements of other levels on the basis of grouping constraints derived directly or indirectly from domain knowledge.

### 3.3 Types of relationships between locations

#### 3.3.1 The inheritance of Gestalt theory

The emphasis on relationships rather than autonomous parts is not novel. It has strong similarities to a line of thought which evolves from Gestalt psychology and which suggests that the relations between parts and wholes are the following: “(1) Wholes are constructed from the parts in the proximal stimulus by a process of description of how the parts relate to one another. (2) Wholes have emergent properties based on these relationships that, by definition, do not inhere in the parts. (3) Wholes have natural, phenomenal subparts that constitute an important aspect of their appearance, but not all wholes have such phenomenal parts. (4) Natural parts, where they exist, need not necessarily be the same as the parts from which the whole is constructed (for example, in drawings of three-dimensional figures, corners are important building blocks of the whole but are not necessarily perceived as subparts). (5) Wholes modify or even camouflage how parts appear, and this fact bears on 3 and 4. (6) Related to 5 is the fact that the description of wholes may be global, in which details and nuances drop out. (7) The final description has attributes or features that characterize the whole, and thus these features do not inhere in the parts from which the whole is constructed. Therefore, a listing of such attributes should not be confused with a list of features that do inhere in the parts” [Rock 1983, p.86].

The present subsection is an introduction to the Gestalt principles of organization as related to grouping in  $\text{CHQP}$ . These principles should be considered with respect to the objective of grouping, which in Gestalt theory is normally the identification of *entities*, as opposed to the identification of *relationships* in  $\text{CHQP}$  where locations, the atomic parts of an entity, retain their integrity throughout all stages of grouping, something that runs contrary to the anti-elementarism of Gestalt theory.

The Gestalt principles of organization have had a significant and lasting influence on computer vision and in particular to systems concerned with texture. Given that texture is of paramount importance in real-world images, it is not uncommon to find Gestaltist elements even in approaches which relate more to psychophysics [Marr 1982]. This is hardly surprising if we consider the state of the art in the psychology of perception. Despite the general decline of hard-core Gestalt theory, “as far as ... the Gestalt laws of groupings [are concerned], little progress has been made in the last half century. These principles continue to be illustrated in textbooks, and except for some work on grouping by similarity in texture patterns ..., and the uncovering of a principle of preference for convexity in figure-ground organization ..., no new ‘law’ has been added, nor have any original ‘laws’ been shown to be incorrect” [Rock 1983, p.74]. The Gestalt principles of organization

still seem to be the only powerful generalizations about the perception of patterns [Hochberg 1972, p.72]. In addition, the emphasis on spatial organization and the use of abstract two-dimensional images, which have much in common with abstracted binary images of architectural plans, makes Gestalt theory an obvious starting point in the investigation of types of group formation in CHQP.

The relevance of Gestalt principles of organization does not imply that the whole Gestalt system of spatial organization is applicable in CHQP. The reason is that, although the principles can be accepted independently from a variety of viewpoints, Gestalt theory in general poses several problems that defeat its universal validity. For one, Gestalt theory involves a particular aesthetic bias, as revealed by the associations of *Prägnanz*, the general principle which determines the best possible psychological organization of stimuli under the prevailing conditions, with unity, uniformity, good continuation, simplicity of shape and perfect organization; and the association of closed areas and good shape with stability (“A ring is a perfectly balanced figure with no articulation within it” [Koffka 1935, p.135]). For Gestaltists ‘good’ embraces regularity, symmetry, simplicity, etc. [Koffka 1935, p.110].

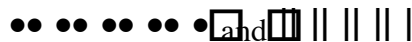
Such biases cannot be justified in terms of information-theoretic or other economy [Attneave 1954; Buffart et al 1981; Garner 1974; Garner & Clement 1963; Hochberg & McAlister 1953; Leeuwenberg 1971, 1978; Palmer 1982, 1983] for two related reasons. The first is that principles of spatial organization apply not to the cortical-retinal stimulus array but rather to the *perceived* image elements [Rock 1983, p.75]. For example, the perceived distance between two points is not that between their retinal projections but the estimated distance of these points in the real three-dimensional world. Therefore, it is unlikely that measurements of economy on the retinal image and on the finally perceived scene description would produce the same value. Also, if a retinal image can be perceived as two different, equally acceptable real-world scenes (not facets of the same scene, as the Necker cube) then the same retinal image should support two different economy values, one for each description. Measurement of economy on the basis of the retinal image results, however, into a single value.

The second reason is that measurement of the economy of each alternative organization presupposes that all alternatives are clearly recognized and somehow encoded, a function which is unlikely in human perception, given the very high number of alternatives [Attneave 1982; Rock 1983, p.163]. As we have seen in the discussion of adapted chain coding in subsection 3.2.1, the number of alternative groups a location may belong to is quite high, despite the restrictions imposed by the orthogonality constraints and the

limited number of primitive group types.

In conclusion, a standpoint that can be safely adopted is that “grouping on the basis of factors such as proximity, similarity, common fate, and good continuation is the result of a decision to describe the stimulus array in one way rather than other possible ways” [Rock 1983, p.75]. From this standpoint the following Gestalt principles of spatial organization are considered as criteria of preference for particular grouping relationships between locations:

Proximity and equality. In fields with ‘equal’ parts proximity easily determines group formation. For example, we see both



as forming pairs [Koffka 1935, p.164]. In fields with different kinds of atomic units, equality (of shape, size and colour) is for Gestaltists stronger than proximity, which in turn is more significant than equality of colour [Koffka 1935, p.165; Wertheimer 1938, p.72]. Figure 3-22 depicts a sequence of examples which support this preference order. In Figure 3-22a bars of different shade (colour in the original experiments) form groups on the basis of proximity. In Figure 3-22b pairs are formed by bars and wavy lines on the basis of proximity, although preference over pairs formed by bars and wavy lines only is only slight. In Figure 3-22c pairs are formed by bars or wavy lines only, bars and wavy lines having different shades.



**Figure 3-22** Proximity and equality

*left (a): proximity is more significant than equality of colour: pairs are formed by bars of different shade; middle (b): equality of shape is approximately as significant as proximity: pairs are formed either by two bars or by two wavy lines; right (c): the addition of colour makes it clear that equality is more significant than proximity: pairs are formed by wavy lines or by bars only (after Koffka [1935])*

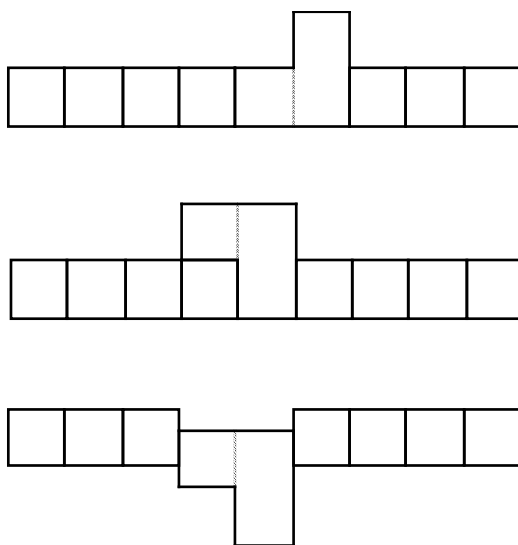
The relevance of proximity and equality to grouping in the initial version of CHAP is rather restricted because, of the factors governing equality, proportions of locations are generally not taken into consideration in the grouping process. The reason for not including proportions at this stage of CHAP is that they relate mostly to aesthetic appreciation or ‘figural goodness’. Following the subdivision of modules in CHAP, this ‘goodness’ depends upon the conformity of the location, the location group or the group configuration to the formal constraints of the classical canon, which is the topic of Chapter 4.

In this chapter most examples are of architectural plans

comprising only rectangular locations and therefore there will be little opportunity to show the significance of shape in grouping. In general, complex shapes are decomposed with respect to the relationship considered against. If all components that result from decomposition support the relationship then difference in shape does not disrupt the relationship (Figure 3-23 — also Figure 3-29).

In grouping rectangular locations shape (aspect ratios) and size do not play a significant role. Locations of different shape and size can be grouped together if they support the same relationships. Size enters only indirectly: the larger a location, the more contiguous locations it may have and therefore the more grouping relations it may assume with different locations.

Proximity at the group recognition stage in chap generally assumes a very narrow sense, that of *contiguity*. Two locations are either contiguous or not; a grouping relationship is normally acceptable only if it occurs among successively contiguous locations, that is, when the relationships are continuous and uninterrupted. On higher levels allowances are made for interruptions in otherwise homogeneous groups (Figures 3-23b, c) but this is normally applies to relations between recognized groups and not among locations.

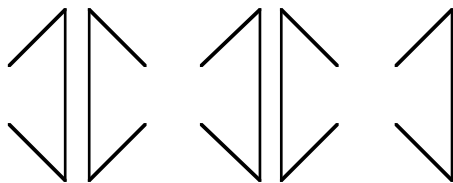


**Figure 3-23** Location shape and decomposition in grouping  
top (a): the L-shaped location does not disrupt the continuity of the alignment group because both its components are aligned; middle (b): the L-shaped location does not form an alignment group with the others because, although the alignment front is not disrupted, only one of the components of the L-shaped location is aligned ; bottom (c): the L-shaped location disrupts the alignment group because it disrupts the alignment front

Closure. Closed organization is considered more significant than proximity and good continuation [Koffka 1935, p.168]. In Figure 3-24 pairs are formed by opposite facing brackets, despite the proximity of back-to-back

brackets. In Figure 3-25a we see two continuous intersecting lines, while in Figure 3-25b organization in two triangles is preferred over the good continuation factor of the two intersecting lines.

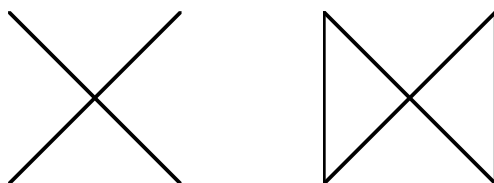
Criteria related to closure are used on the higher levels of group formation in CHQP, that is, in relating recognized groups of locations into larger and more complex configurations, yet equally coherent as their constituent groups. In these cases one might say that closure is stronger than proximity, proximity being contiguity among groups. Such mergers are instrumental in resolving ambiguities concerning which group a location belongs to (Figure 3-26).



**Figure 3-24** *Closure versus proximity*  
pairs are formed by two facing brackets, although the distance between back-to-back pairs is shorter (after Köhler [1929])

Good shape. The circle is a frequent example of good shape in Gestalt theory. Good shape is determined by the coincidence of expectations formed on the basis of discrete parts of a shape with the whole shape: “Each piece of [a good shape] contains the principle of the whole” [Koffka 1935, p.151].

Goodness of shape does not enter group formation in CHQP. Goodness in general is considered an element of cognitive filtering, which is discussed in Chapter 4. In the framework of cognitive filtering, goodness of shape can be judged in terms of the proportions of a location. Despite the frequent occurrence of regular arrangements in classical architecture (for instance arrangement of locations in the shape of a rectangle around an atrium, as in Figure 3-26), such regularity is treated as an issue of homogeneity between the relationships of the groups which form the overall arrangement (i.e., more related to closure and continuation — also Figure 3-28) than to goodness of shape.

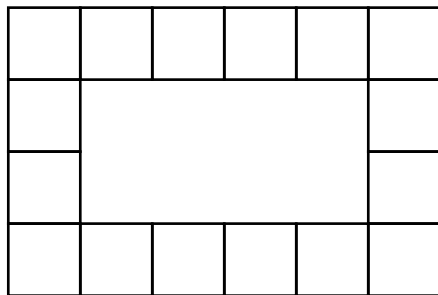


**Figure 3-25** *Good continuation*  
left (a): the image is perceived as comprising two units, two intersecting lines; right (b): the addition of two more lines causes reorganization of groups: the image is now perceived as two touching triangles (after Koffka [1935])

Good continuation. Good continuation (Figure 3-25a)

can be described on the basis of uniformity of repetition. Good continuation is closely related to good shape and stability: “a straight line is a more stable structure than a broken one, and ... therefore organization will, *ceteris paribus*, occur in such a way that a straight line will continue as a straight line. We may generalize thus: any curve will proceed in its own natural way, a circle as a circle, an ellipse as an ellipse and so forth” [Koffka 1935, p.153]. Figure 3-27 is a typical case where “the external forces prevent good organization” [Koffka 1935, p.153].

Continuation is of particular concern to *chop*, in the sense that groups of locations are characterized by continuous relationships. It also enters higher levels of group formation, in the form of justification for allowances for interruptions to otherwise continuous relationships, as in Figure 3-23c. In these cases continuation is stronger than proximity, similarly as in Figure 3-28.



**Figure 3-26** Merger of simple (first level) groups into complex (second level) groups

*the image is initially described in terms of four alignment groups, with the corner locations being members of two groups each; the relationships between alignment fronts of different groups allow the connection of the four groups into one, thereby resolving the ambiguity concerning the corner locations*

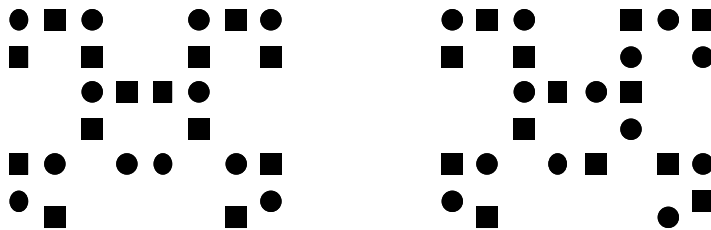
Gestalt theory has been criticized on a number of issues, such as the aforementioned aesthetic bias, including the reduction of ‘good figure’ or ‘closure’ to approximations of specific shapes, such as circles [Gregory 1970, p.20]. On a different level, Gestaltist have been criticized for the selection of abstract images as examples of the visual world: “... the abstract space of points, lines and planes [is] a poor conception with which to begin the analysis of how we see, for no one has ever seen it” [Gibson 1950, p.188]. Also, although “the ‘Principles’ were developed by inductive generalization from instances” on the basis that “since we all experience essentially similar objects it would not be too surprising if we developed similar, even identical, generalizations” [Gregory 1970, p.21], it is impossible to find general agreement in a universal preference order of grouping principles.



**Figure 3-27** *Good continuation hindered by external forces*  
 “The result is an esthetically unpleasant impression, because the proper continuation of the four semicircles is interrupted” [Koffka 1935, p.153]  
 (after Bühler [1913])

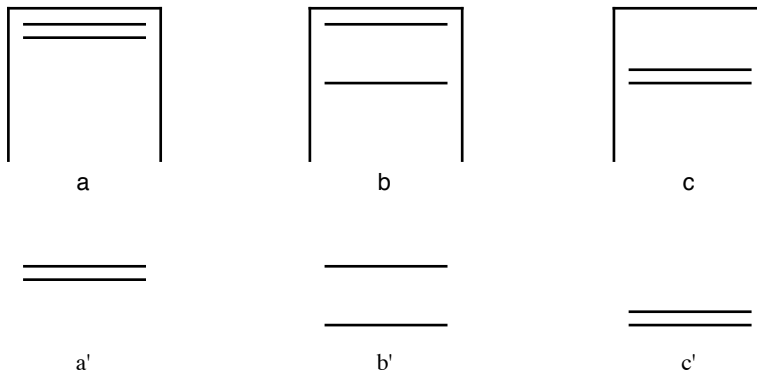


**Figure 3-28** *Alignment versus proximity*  
 (with reference to the three bullets where the two perceived lines appear to meet) “Evidently the tendency to organize dots into rows is stronger than the tendency of association by proximity” [Gregory 1970, p.20].



**Figure 3-29** *Similarity based on overall articulation*  
 the two figures appear to be identical unless they are examined carefully:  
 half of the homologous parts are in fact different (after Palmer [1982])

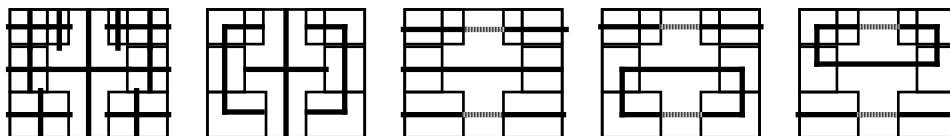
Most of this criticism does not apply to chap because chap deals with abstract line drawings and not real-world images and also because many of the Gestaltist biases are avoided thanks to scaling-down. Furthermore, recognition of the spatial articulation of an architectural plan refers more to general (multilateral) relationships which in fact are combinations of different grouping criteria. Within these relationships each criterion plays a different role and is usually very closely connected to other criteria. A parallel to that can be observed in Figure 3-29 where the overall articulation of the stimulus and not a particular criterion dominates. Also, in a modular system with several levels of grouping each criterion has different significance on each level and for the groups which result from each level, similarly to Figure 3-30.



**Figure 3-30** *Preferences on different levels of grouping*  
*b is generally considered more similar to a than c although c' is obviously more similar to a' than b' (after Goldmeier [1972])*

### 3.3.2 On the applicability of generalized cones and medial axis transforms

The structure of the recognition and representation of location groups suggests that representations based on configurations of component axes, such as generalized cones and medial axis transforms, might be applicable to the grouping stage of chap. In the previous chapter (subsection 2.5.2) some of the general disadvantages of these techniques were discussed with respect to their application to orthogonal architectural plans. These include a general lack of distinction between different repetition patterns in asymmetrical shapes (Figure 2-24) and overstatement of axial symmetries (for which they were actually developed). In this chapter we assume an advanced adaptation of the generalized cones technique, based on the recognition system by Nevatia & Binford [1977], on the decomposition techniques by Pavlidis [1977, chapter 9] and on the smoothed local symmetries by Brady [1982, 1983]. This adaptation accepts more than one axes for each location and also assumes that the configuration of the axes contains all information about the structure of the architectural plan.



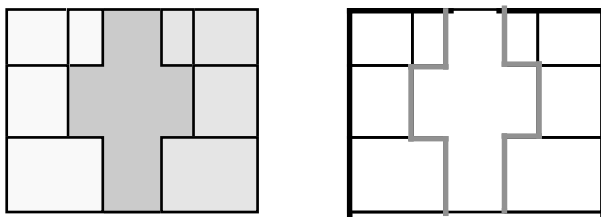
**Figure 3-31** *Description of location groups by configurations of component axes*

*from left to right: the axes of each location; alternative configurations of axes which correspond to different groups of locations (in order of preference); broken lines denote parts of group axes that do not correspond to location axes*

The use of medial axes to describe the spatial arrangement of locations is adequate with respect to only

some of the spatial relationships between locations (Figure 3-31). With respect to the three relationships of alignment, transposition and alternation introduced in the previous chapter (subsection 2.5.4), description of location groups by the configuration of their medial axes appears to be better suited to cases of alignment on both sides as these support better the underlying concept of axis continuation. This has been recognized as a potential problem in computer vision and it has been suggested that there should be allowances for cases where the axis of a shape is part of the bounding contour [Brady 1983, p.49 & Fig.11], which in our case corresponds to the alignment front.

As far as relations between different groups are concerned, medial axes are helpful in distinguishing the various major subdivision of a plan —bipartite and tripartite subdivision into zones in the case of Figure 3-31. Relations between the different groups which correspond to these subdivisions are less clearly described due to the detachment of the axes from the boundary where these relations are more obvious.



**Figure 3-32** *Description of location groups by relationships of alignment, alternation and transposition*

*left: the three groups of locations recognized in Villa Foscari; right: relationships of alignment (thick black lines) and alternation (gray lines) in each group*

To anticipate section 3.4, a description of the same architectural plan in terms of alignment, transposition and alternation groups integrates more of the aspects that concern an architect (Figure 3-32). The orientation of the groups which so prominent in medial axis configurations is described by the orientation of the alignment and alternation fronts which are far less sensitive to aspect ratios and shape perturbations than medial axes. In Figure 3-31 this sensitivity was reduced effectively by the techniques of smoothed local symmetries and local frame propagation [Brady 1982, 1983] but this an additional computational burden not required by the grouping techniques of chap.

A comparison of Figure 3-31 to 3-32 reveals the fundamental differences between the two representations. The representation derived from generalized cones is essentially a *shape* representation with little concern for the number of entities (i.e., locations) each shape comprises. All that matters is the overall shape of each group. This means that grouping is based on shape constraints, that is, on the acceptability of a common medial axis for the whole group.

As a result, the number of alternative group configurations is quite high even for the scaled-down environment of chap (Figure 3-31).

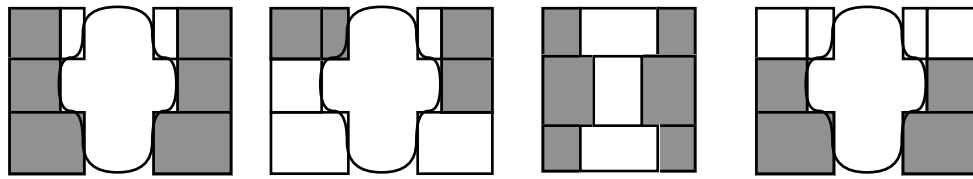
In the representation of Figure 3-32 grouping relies on *structural* relationships between locations, that is, relationships which essentially amount to underlying elements in the architectural plan, such as load-bearing and circulation patterns. Recognition of these relationships is more sensitive to contiguity conditions than to changes in shape and size so long as contiguity is not affected. Grouping relationships are identified in particular parts of the locations: an alignment front lies on a specific side of all locations aligned to it. This facilitates (a) recognition of the underlying physical elements or design decisions that resulted into the particular grouping relationship through the recognition of the regularities they have caused, and (b) correlation of groups related by type and position, that is, of groups that have resulted from the same or related underlying physical elements or design decisions.

In conclusion, a proper adaptation of generalized cones and medial axis transforms, similar to the one outlined in the present section, could in principle be applied to the description of orthogonal architectural plans. However, it seems that such an adaptation is not ultimately capable of differentiating between the various relationships between locations and groups, nor to the subsequent multiplicity of interpretations. Although such aspects are implicitly described by medial axis configurations, these configurations seem to be too abstract to cover the whole spectrum of relationships and all levels of description, mainly because they have been developed for the description of forms with strong physical medial axes, while in architectural plans grouping relationships are particularly evident on the boundaries of locations, as also in texture discrimination [Muller 1986], where generalized cones and medial axis transforms are not widely used.

### 3.3.3 Multilateral relationships between locations in chap

The relationships that concern chap are essentially those which characterize location groups in adapted chain coding (Figure 3-7) with the improvements suggested in Figures 3-17 and 3-18. As in adapted chain coding the relationships and resulting groups are kept as simple as possible and more complex groups are treated as aggregations or fusion of such simple groups. Also as in adapted chain coding a relationship may have either vertical or horizontal direction, although

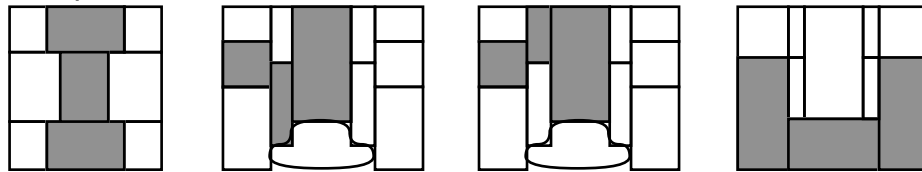
relationships and groups are not labelled with their orientation.



**Figure 3-33** *Examples of alignment relationships in architectural plans*

*from left to right: groups with a single alignment front; groups with two alignment fronts; aligned groups also characterized by alternation; aligned groups also characterized by transposition*

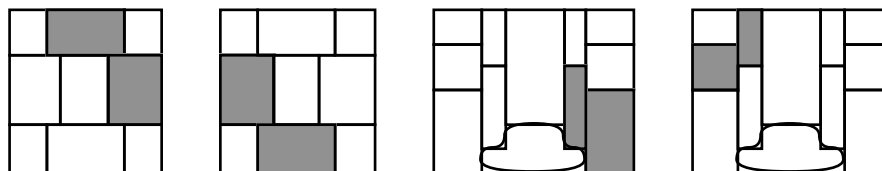
The multilateral relationships that correspond to the bilateral ones we have considered in adapted chain coding (Figure 3-6) are exactly those used for the identification of location shape type in Chapter 2 (Figure 2-29 to 2-31): alignment, alternation and transposition.



**Figure 3-34** *Examples of alternation relationships in architectural plans*

*from left to right: two groups with two alternation relationships and two groups with a single alternation relationship*

*Alignment* groups (Figure 3-33) are characterized by one or two *alignment fronts*, that is, cover adapted chain coding group types 10, 11/13, 12/14, 20, 21/23 and 22/24, as well as fusions and aggregations thereof in the manner of figure 3-17. Alignment groups are of two basic types: groups with a single alignment front and groups with two alignment fronts. The former may also be characterized by a relationship of a different type (i.e., alternation or repetition) on the side opposite to that of the alignment front.



**Figure 3-35** *Examples of transposition relationships in architectural plans*

*groups with two transposition relationships*

*Alternation* groups may also have two alternation relationships, a single alternation relationship, or an alternation and a transposition relationship (Figure 3-34). In the case of two alternation relationships these may have the

same direction and pattern or not, exactly as in subsection 2.5.4. Transposition groups may have two transposition relationships of the same or different direction and pattern or a single transposition relationship (Figure 3-35).

These types are the relationships which concern the first level of grouping in CHQP. As in adapted chain coding, a preference order determines the precedence of each relationship and each group type. The preference for grouping relationships is the same as in adapted chain coding:

- 1 alignment
- 2 alternation
- 3 transposition

For group types the preference order deviates slightly from that of adapted chain coding group types and is as follows:

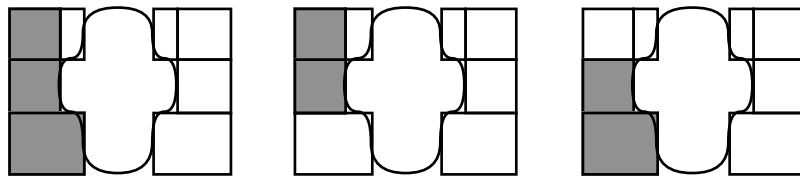
- 1 groups with two alignment fronts
- 2 groups with an alignment front and an alternation relationship
- 3 groups with an alignment front and a transposition relationship
- 4 groups with two alternation relationships
- 5 groups with an alignment front only
- 6 groups with an alternation and a transposition relationship
- 7 groups with two transposition relationships
- 8 groups with a single alternation relationship only
- 9 groups with a single transposition relationship only

This order reflects formal, functional and constructional constraints and spatial articulation types of classical architecture. A complete justification goes beyond the scope of CHQP as it would require correlation of a number of disparate sources of evidence, such as the predominance of linear articulation in classical architecture [Tzonis & Lefaivre 1986, chapter 5] and preference for linear over other types of arrangement in human vision [Attneave 1982; Perkins 1982]. Instead, the group types in CHQP and their preference order are verified by the comparison of the results of grouping with our intuitive interpretations of an architectural plan and therefore should be accepted as a faithful even if not completely accurate and precise representation of the lower grouping constraints in classical architectural plans. Any alteration to the preference order or the group typology would not require any structural modifications in the grouping processes of CHQP and would moreover cause little change to the resulting descriptions of architectural plans as the primitive group types we have considered so far are combined into more complex arrangements on subsequent grouping levels. Furthermore, CHQP concerns only the geometry of architectural plans and that on a rather high level of abstraction. Additionally, for instance functional, viewpoints provide further grouping

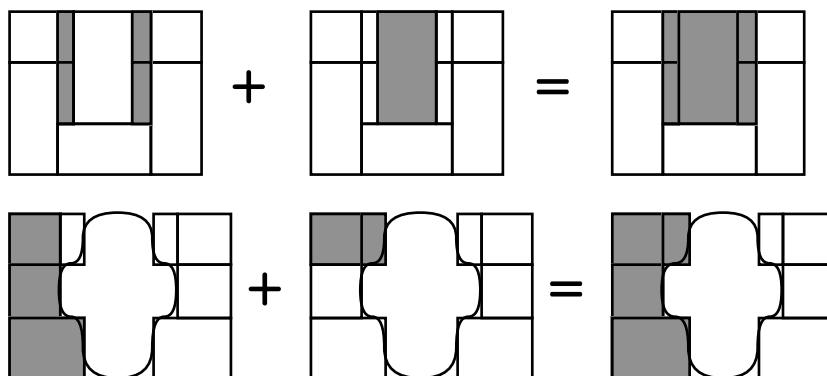
constraints that may improve grouping accuracy and precision.

So far we have discussed the grouping of rectangular locations only. Inclusion of orthogonal location of complex shape follows the same principles and uses the same techniques as in adapted chain coding: locations of complex shape are decomposed into rectangular slices and grouping is then considered in terms of such components and rectangular locations. As is adapted chain coding all components of a location must belong to the same group and for a location to belong to a specific group of this level all its components must be acceptable as members of the group as if they were independent locations. These rather stringent constraints are relaxed on subsequent levels of grouping so as to allow for local or partial deviations and interruptions.

A particular point about the decomposition of locations of complex shapes is that the orientation of the rectangular slices is not necessarily horizontal, as in the Chapter 2 (subsection 2.5.3). Instead, as in adapted chain coding, complex shapes are decomposed into either horizontal or vertical slices, depending on the orientation of the relationships under investigation. For example, when investigating vertical alignment fronts, locations of complex shapes are decomposed into horizontal slices.



**Figure 3-36** *Decomposability of simple (first level) groups*  
*a group (left) may be decomposable into two or more groups (middle and right)*  
*but these are characterized by a common relationship that offers no reasons for*  
*decomposition*



**Figure 3-37** *Examples of complex (second level) groups*  
*top (a): on the basis of proximity; bottom (b): on the basis of closure*

The group types presented so far cover the *first level* of grouping in chcp. This concerns the recognition of *simple* groups characterized by a single direction of articulation and

hence by one or two instances of alignment, alternation or transposition. Simple (first level) groups cannot be decomposed into subgroups that do not share a common uninterrupted relationship. For example, either vertical alignment group in Villa Foscari can in principle be decomposed into two overlapping alignment groups (Figure 3-36). These, however, share a common alignment relationship which offers no reason for decomposition (no lacuna and no change of direction).

This first level is followed by additional grouping procedures which form the *second level* of grouping. On this second level, groups recognized on the first level are combined into *complex* groups which generally correspond to conventional architectural articulation types. In view of the lack of extensive typologic studies of classical architectural plans and of their parts which would have provided the typologies that should concern visual architectural recognition, the types of complex (second level) groups in CHQP has been kept to the absolute minimum that can be produced by the recursive application of a limited number of combinatorial operations on simple (first level) groups or rather on their respective grouping relationships. These operations are:

- Aggregation of groups with (mostly solitary) locations that are contiguous to all members of the group on sides that are characterized by a grouping relationship. This operation results into location patterns generally associated with circulation types, such as single or double loaded corridor arrangements (Figure 3-37a).
- Fusion of intersecting similar groups. This operation allows the fusion of groups that have common members and intersecting relationships of the same basic type. The relationships may have the same or different orientation. In the former case the relationships have necessarily different patterns of repetition (such as, in the terminology of the previous chapter, descending and ascending transposition). In the latter case it is the type of intersection that counts. For example, two alignment groups with fronts that meet at right angles and have a common endpoint are normally combined into an L-shaped alignment group (Figure 3-37b).
- Aggregation of adjacent groups characterized by exactly the same relationships in the same direction of articulation. This operation essentially amounts to bridging lacunae that may occur in a group. This operation is the most infrequent of the three because it requires that the location(s) that occupy the lacuna are also integrated into the complex group, that is, that their relations to other locations are not strong enough to form a group.

## 3.4 Recognition of grouping relationships

Group formation in chap is a bottom-up bilevel process which proceeds from the recognition of simple multilateral relationships between locations of an architectural plan to the recursive combination of the groups defined by these relationships to the description of the plan as a configuration of mutually exclusive groups of the highest possible complexity.

The first of the two levels concerns the recognition of simple (first level) groups, that is, groups characterized by one or two relationships between all their members and a single direction of articulation. The output of the first level is a set of partially overlapping simple groups. The second level combines the simple groups into complex (second level) ones and determines the compatibility of the resulting simple and complex groups. The output of the second level is one or more alternative descriptions of the plan as a configuration of mutually exclusive groups of locations.

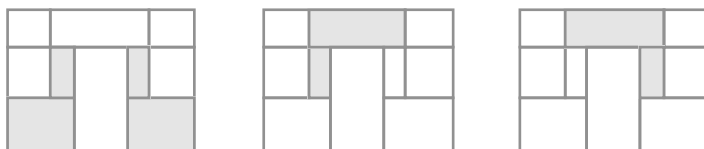
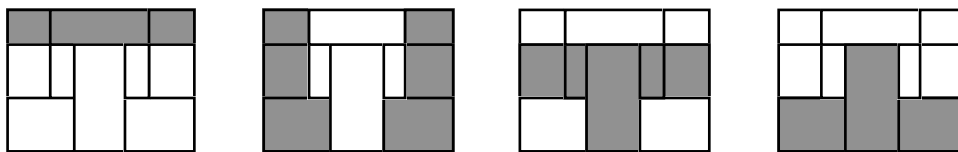
### 3.4.1 The first level of group recognition: alignment, transposition and alternation

The first level of grouping in chap concerns recognition of alignment, alternation and transposition relationships between groups of locations. In a serial system recognition of these relationships follows temporally the preference order of subsection 3.3.3: first alignment is investigated, then alternation and finally transposition. Instances of these relationships that are recognized between specific locations at one of these three stages are binding for subsequent stages in that they restrict investigation of the other relationships to distinct location subsets. The exact order and structure of steps in first level grouping is:

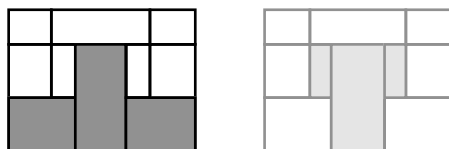
- 1 investigation of alignment relationships
  - 1.1 recognition of simple (first level) alignment groups
- 2 investigation of alternation relationships
  - 2.1 in each alignment group with a single alignment front and
  - 2.2 among the remaining locations of the plan (i.e., those

- that do not belong to already recognized groups)
- 2.3 recognition of simple (first level) alternation groups and of alignment groups also characterized by alternation
- 3 investigation of transposition relationships
    - 3.1 in each alignment group with a single alignment front
    - 3.2 in each group characterized by a single alternation relationship and
    - 3.3 among the remaining locations of the plan (i.e., those that do not belong to already recognized groups)
    - 3.4 recognition of simple (first level) transposition groups and of alignment or alternation groups also characterized by transposition

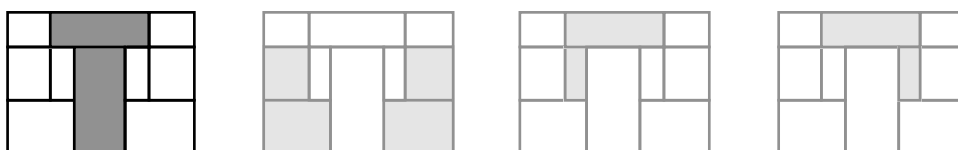
As indicated in Figure 3-38, at each stage (that is, in the investigation of each relationship type) overlapping groups are recognized. Some of these are immediately rejected for any of the following three reasons: (a) their locations form larger or stronger (i.e., with more relationships) groups with other locations at the particular stage; (b) they are subsets of groups identified at previous stages; (c) they have already been rejected for either of the above reasons at a previous stage.



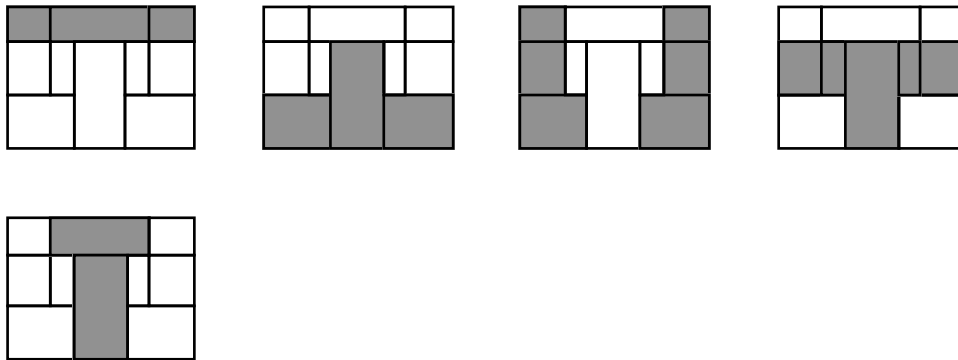
**a. alignment:** top: accepted groups; bottom: rejected groups (because their locations form larger or stronger (i.e., with two fronts) groups with other locations)



**b. alternation:** left: accepted group; right: rejected group (because it is a subgroup of an already accepted group)



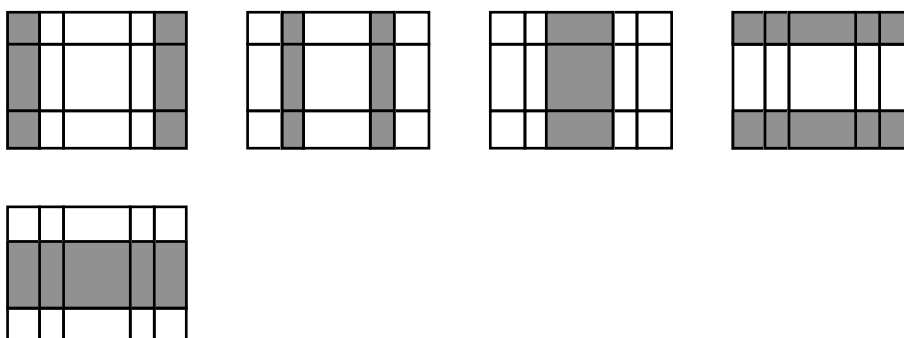
**c. transposition** (from left to right): accepted group; two groups rejected because they are subgroups of already accepted groups; two groups rejected because they were already rejected in the investigation of other relationship types



*d. simple (first level) groups*

**Figure 3-38** *First level of grouping in the case of Palladio's Villa Zeno*

Investigation of each relationship type locally, that is for pairs of locations, essentially relies on adapted chain coding. Still, an attempt is made to utilize the multilateral character of the relationships. This is more easily accomplished in the investigation of alignment, as what appears at first sight to be an alignment front is always an alignment front, even if the group it defines is rejected (i.e., not preferred over alternative alignment groups). Recognition of alignment relationships begins with the identification of potential alignment fronts in the plan, that is, edges which form a continuous straight line segment. Then locations which are aligned on each front are identified. These form two distinct sets, depending on whether there are to the left or the right of, above or below the front. Within either set continuous groups of more than two locations are accepted as alignment groups. The final step in the recognition of alignment groups is the identification of groups which have two alignment fronts, that is, of cases where exactly the same locations are aligned on two distinct fronts. If just some of the locations of an alignment group have a second front, then this second front is ignored on the first level of grouping.

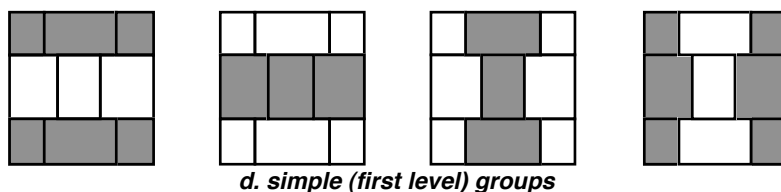
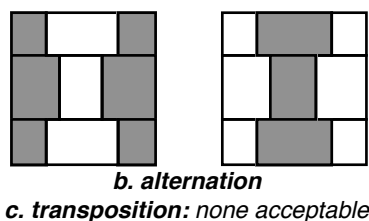
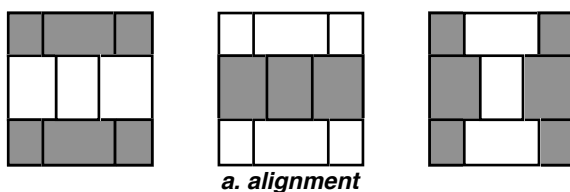


*a. alignment*  
*b. alternation: none*  
*c. transposition: none*

*d. simple (first level) groups: as in (a)*

**Figure 3-39** First level of grouping in the case of Palladio's Villa Thiene

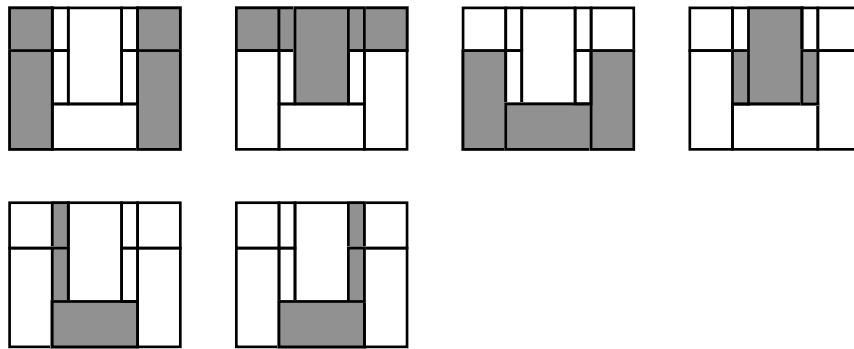
Investigation of the two other relationship types does not follow the same procedure because alternation and transposition patterns among edges do not necessarily correspond to alternation and transposition patterns among locations. Alternation and transposition are recognized through a variation of adapted chain coding which employs expectations. The expectations are (a) general, i.e., focus search only on the appropriate kinds of contiguity relations, and (b) particular, i.e., formed on the basis of previously recognized contiguity relations. For example, if the relationship between two contiguous locations suggests the possibility of an alternation group, we can have concrete expectations about the direction of articulation of this group and its alternation type. These expectations determine the position of and contiguity relationships between potential new members of the group. Locations which occupy the corresponding positions are checked for this particular alternation type and group. Acceptance or rejection of alternation and transposition groups is the same as for alignment groups.



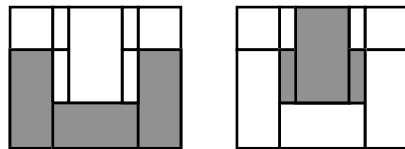
**Figure 3-40** First level of grouping in the case of Palladio's Villa Ragona

An issue open to question is whether transposition and alternation should be treated as variations of alignment, i.e., alignment along oblique or discontinuous fronts, that

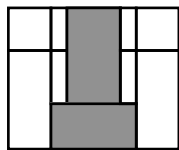
produce classes of groups which, although different in appearance, are similar in organization to alignment groups. Answers to this question can only be given on the basis of extensive analysis of architectural plans through a proper representation system. Therefore CHQP becomes also a prerequisite for the resolution of the issue. As all computer systems, CHQP requires an initial period of trial and training during which representations and procedures are fine-tuned so as to achieve better coordination and more precise correspondence to the application domain. Only then it is possible to discuss such issues with high degree of certainty. For the moment it is sufficient to state that even if transposition and alternation are just different kinds of alignment, the mere deviation of their fronts from orthogonality and/or continuity constraints implies the necessity of special treatment.



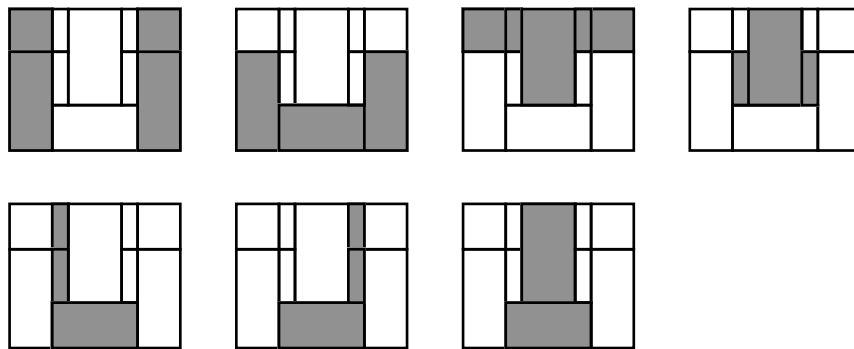
*a. alignment*



*b. alternation*



*c. transposition*

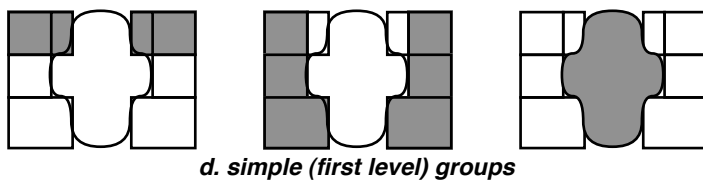
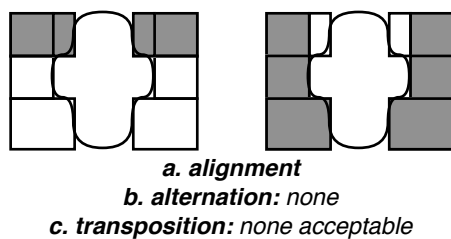


*d. simple (first level) groups*

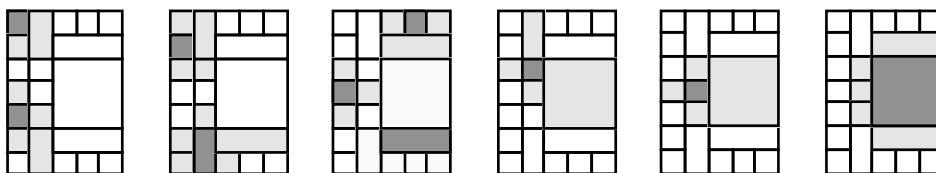
**Figure 3-41**

*First level of grouping in the case of Palladio's Villa Poiana*

A device optionally used is the *neighbourhood* of a location, that is, the set of locations that are contiguous to it (Figure 3-43). Location neighbourhoods are used in many ways: the size and position of the largest neighbourhoods in an architectural plan often reveals aspects of its organization, while in large or complex architectural plans (i.e., with scattered alignment, transposition and alternation groups) it is useful to subdivide the subsequent large set of locations into smaller subgroups which can be treated more efficiently. Investigation of alignment, alternation and transposition relations initially within location neighbourhoods and then extension and merger of recognized groups may increase efficiency even for moderately large and complex architectural plans.



**Figure 3-42** First level of grouping in the case of Palladio's Villa Foscari



**Figure 3-43** Location neighbourhoods in an architectural plan  
light gray locations represent the neighbourhood of the dark gray location they surround (not all neighbourhoods are shown; those which are exactly symmetrical to one of the above with respect to a central horizontal axis are omitted)

The first level of grouping in an architectural plan results into the recognition of a number of partially overlapping groups that are characterized by one or two relationships. Few if any of the locations of an architectural plan remain solitary, i.e., not included in any group. Solitary locations of complex shape are treated as groups in that they are considered as alignment, alternation or transposition groups if their shape type (cfr. Chapter 2) justifies any of these relationships. For example, the central cross-shaped location of Villa Foscari is considered to be characterized by two alternation relationships (Figure 3-42).

The relationships between locations that are recognized

on this level of *chap* form the sufficient and necessary basis of the description of all subsequent levels, in the sense that no further information need be extracted directly from the image of the plan. By manipulating the relationships recognized on the first level of grouping we can identify successively more complex groups up to the level of the geometry of the whole architectural plan. References to the set of locations of the architectural plan or to the plan itself are quite rare and relate mostly to elements and aspects observed on the first level of grouping. This suggests that the design constraints which correspond to these relationships are both fundamental and rudimentary.

A closer examination of Figures 3-38 to 3-42 in juxtaposition to the actual architectural plan reveals that grouping relationships appear to be related to constraints determined by the site, the construction and essential aspects of the functional types used. For example, the existence of doubly aligned groups along the perimeter of the plan is obviously related to the geometry of the construction method used and the topography of the site. Also, the neighbourhoods of locations used for circulation or other public / communal activities are larger than those of other locations. In a sense, the relationships recognized on this level reproduce the first sketches of a solution: the basic construction parameters, the major subdivisions of space, the primary foci of spatial articulation and other aspects of the building that are decided in rough lines at the early stages of architectural design.

### 3.4.2 The second level of group recognition: complex location groups

The first level of grouping in *chap* (Figures 3-38 to 3-42) produces a description of an architectural plan as a set of groups with a single direction of articulation and one or two relationships of alignment, alternation or transposition. These simple (first level) groups are overlapping to a degree such that they cannot be normally combined into a small number of alternative group configurations that describe the plan in a coherent manner. The second level of grouping in *chap* concerns both the aggregation or fusion of simple (first level) groups into complex ones that represent more comprehensive spatial articulation patterns and the coordination of the location groups into coherent descriptions of an architectural plan in terms of its parts. These two objectives subdivide second level procedures into two temporally distinct parts, the *grouping* part and the *coordination* part.

The grouping part concerns the recognition of complex (second level) groups. It consists of three distinct channels, the closure, the continuation and the proximity channel (the names being simply an homage to related Gestaltist principles of organization). Each of these channels accommodates one of the three operations used for the combination of simple groups into complex ones (cfr. section 3.3.3). The three channels operate in parallel and recursively in rounds so as to recombine complex groups recognized in previous rounds. In the first round the input to each channel is the set of simple (first level) groups of the plan, while in subsequent rounds the input to each channel consists of (a) the complex groups recognized in any channel in previous rounds and (b) the set of simple groups of the plan minus the ones that have been combined in the same channel in previous rounds.

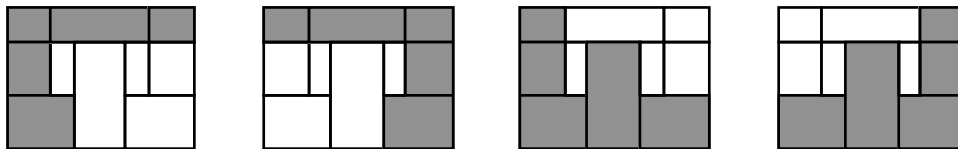
The purpose of the *closure* channel is to recognize patterns of related similar grouping relationships, such as alignment groups that form the perimeter of an atrium configuration (Figure 3-26). In the closure channel the first round concerns the fusion of groups that have common members and intersecting relationships of the same basic type. The relationships may have the same or different orientation. In the former case the relationships have necessarily different patterns of repetition, such as, in the terminology of subsection 2.5.4, descending and ascending transposition. In the latter case it is the type of intersection that counts. For example, two alignment groups with fronts that meet at right angles and have a common endpoint are normally combined into an L-shaped alignment group (Figure 3-44).

In the second round the closure channel first attends to the combination of complex groups recognized in the first round of closure into, for instance, U-shaped alignment groups. Then complex groups of the first round of other channels are considered with respect to each other and to simple (first level) groups that were not combined in the previous round of closure. Subsequent rounds of closure also consider first the groups of the last round of closure and then the products of the last round of other channels with respect to the closure groups of rounds before the last and the simple groups that do not form part of a closure group of higher order.

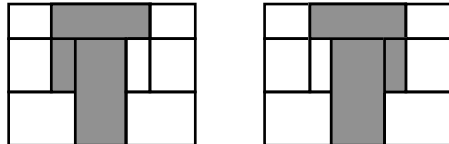
The *continuation* channel aims at the recognition of closely related adjacent groups, that is, groups characterized by exactly the same relationships in the same direction of articulation. This operation essentially amounts to bridging occasional lacunae that may separate essentially the same pattern of repetition of contiguity relationships between locations. The justification of the continuation channel is mainly pragmatic, as such fragmentation of initially

continuous patterns is quite frequent in architectural design, but also relates to the human tendency to encode events in transformed representations which for some reason (usually chunking) are easier to manipulate. For instance, the string *RCH* may be stored as *RICH* which is a meaningful word [Underwood 1972, pp.1–2]. The continuation channel also relates to amodal completion, i.e, completion of a shape whose parts are occluded [Kanizsa & Gerbino 1982]. Nevertheless the continuation channel produces less complex groups than the other two, mainly because it requires that the location(s) that occupy the lacuna are also integrated into the complex group, that is, that their relations to other locations are not strong enough to form a group (Figure 3-48). The operation of the continuation channel is the same as that of the closure channel.

first round

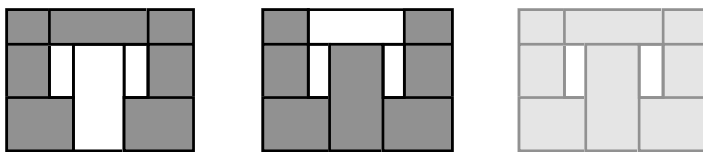


- a. closure**
- b. continuation: none**

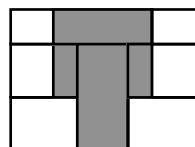


- c. proximity**

second round

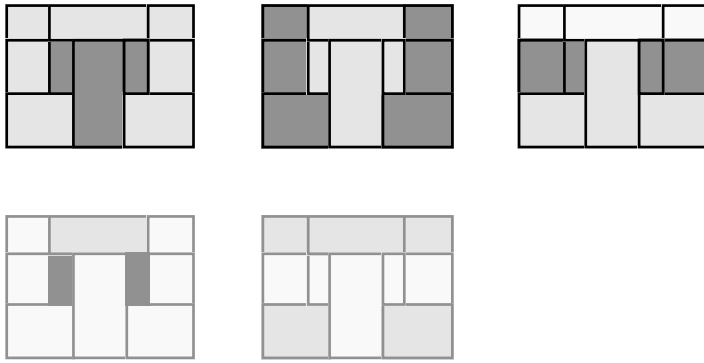


- a. closure: left and middle: accepted complex (second level) groups; right: group rejected because it contains all but two (less than three) locations of the plan**
- b. continuation: none**



- c. proximity**

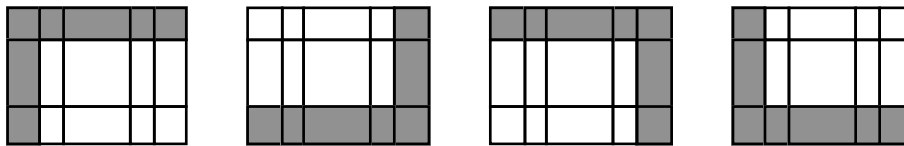
final group configurations



top row: accepted configurations; bottom row: configurations rejected because they decompose groups into solitary locations

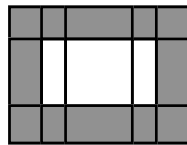
**Figure 3-44** Second level of grouping in the case of Palladio's Villa Zeno (cfr. Figure 3-38)

first round



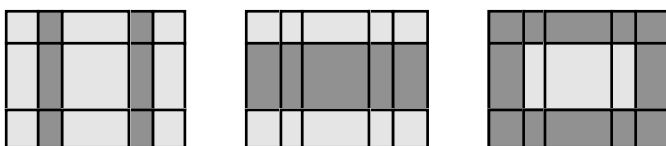
- a. closure
- b. continuation: none
- c. proximity: none

second round



- a. closure

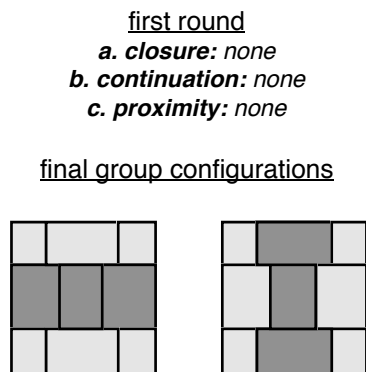
final group configurations



**Figure 3-45** Second level of grouping in the case of Palladio's Villa Thiene (cfr. Figure 3-39)

The third channel, that of *proximity*, relates to the identification of spatial articulation patterns generally associated with circulation types, such as single or double loaded corridor arrangements. In the proximity channel locations that are contiguous to all members of a simple group on sides that are characterized by a grouping relationship unite with the group to form a complex group. The main differences of the proximity channel with the two others are that (a) proximity grouping may require *decomposition* of a simple group so as to extract a location contiguous to another simple group, and that (b) subsequent

rounds of proximity also unite proximity groups of previous rounds that contain the same group or the same solitary location (Figure 3-44). This allows, for instance, the union of two overlapping ‘single loaded corridor’ groups into a ‘double loaded corridor’ group (Figure 3-47).



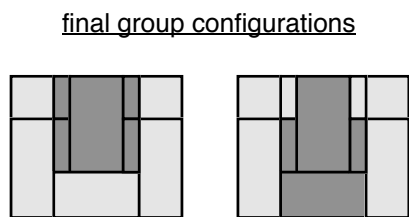
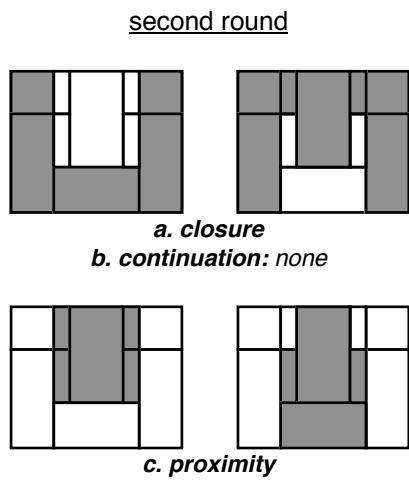
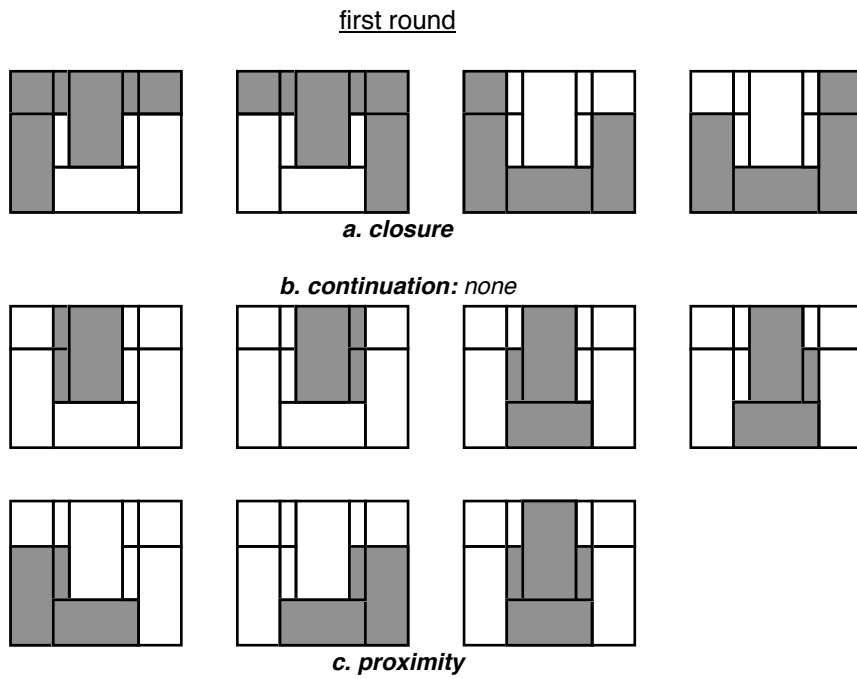
**Figure 3-46** Second level of grouping in the case of Palladio's Villa Ragona (cfr. Figure 3-40)

The recursive grouping activity of the three channels concludes if either of the following conditions is met: (a) No new groups have been recognized in any channel in the same round, or (b) all new groups that have been recognized in a round contain all or all but one or two of the locations of the plan. Such monstrous groups are rejected (Figure 3-44).

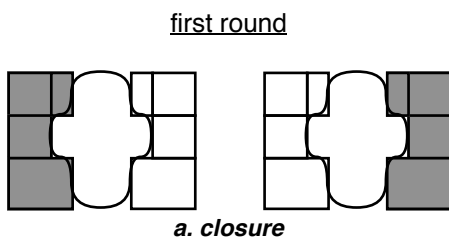
The set of overlapping complex and simple groups returned by the grouping part are organized into coherent configurations of mutually exclusive groups by the *coordination* part. This part first arranges the output of the grouping part in this order or priority: (1) complex groups recognized in the last round of the grouping part, (2) complex groups of previous rounds that form no part of the above, (3) simple groups not contained in complex groups.

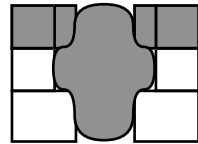
The coordination part investigates first the grouping potential of each of the groups in (1), that is, if each of these groups can be complemented by other groups or fragments thereof so as to form a configuration of mutually exclusive groups that covers the whole plan. The acceptability of resulting configurations is determined by whether they include solitary locations that have resulted from the decomposition of a group (Figure 3-44). An exception is made in the case of solitary locations recognized as such by first level grouping (Figure 3-48). If a complex group forms only rejected configurations, the simple groups it comprises are also considered in the coordination part (Figure 3-44). Other complex groups that may form part of the same rejected group are not considered because they are just intermediate products

of second level grouping.



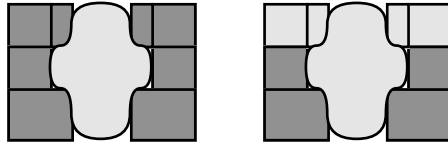
**Figure 3-47** Second level of grouping in the case of Palladio's Villa Poiana (cf. Figure 3-41)





**b. continuation**  
**c. proximity: none**

final group configurations



**Figure 3-48** *Second level of grouping in the case of Palladio's Villa Foscari (cfr. Figure 3-42)*

The second level of group formation is more obviously related to architectural design because it involves recognition of formally and functionally familiar groups of locations, such as single and double loaded corridor arrangements or location arrangements around atria. CHQP does not aim at the recognition of such conventional architectural types because there is no comprehensive typology to refer to. Instead, CHQP is restricted to the description of architectural plans in terms of the spatial relationships of its locations which determine grouping and subdivision. The purpose of the resulting descriptions is to provide the means for the intelligent manipulation of an architectural plan by the computer on levels of abstraction similar or equivalent to those of the computer user. Such manipulation includes classification according to some extrinsic to CHQP typology that can be added to it as a superimposed interface module. Classification is in fact facilitated by the detailed hierarchical analysis and description of spatial relationships in each plan.

### 3.5 Preliminary conclusions

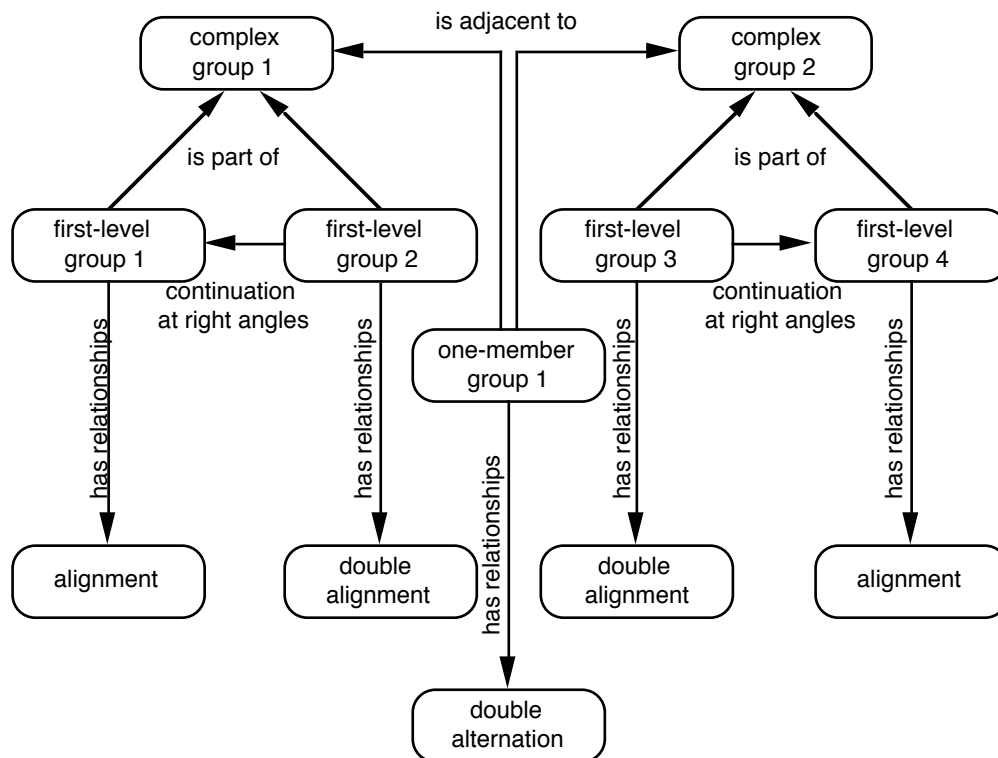
Chapter 3 describes a bottom-up derivation of the description of an image which in general lines complies to current computer vision theory and practice. Any deviation is more due to constraints of the application domain, i.e., architectural plans, than to methodologic differences. What distinguishes CHQP from what could be called 'conventional' recognition systems [e.g. Ettinger 1988] is that the descriptions returned by the grouping process of this chapter are not matched to known prototypical part or whole models so as to recognize these models in the image (in our case the type of the plan or of its component groups).

The absence of such conventional type recognition in CHQP is due the application domain in two ways. The first is

that the lack of typologic studies of classical architectural plans and their subdivision does not allow the definition of group and plan types in CHQP by reference to some established theory of architectural typology. The two options left are (a) to use a rough collection of types as a temporary replacement of a more accurate typology, or (b) to develop CHQP as an open-ended descriptive system which can accommodate any compatible architectural typology. The first option involves the danger of undermining the actual achievements of CHQP by presenting its output in terms of a naive, incomplete or unsatisfactory typology. The second option involves a certain degree of fuzziness in the output of the grouping process which could be a major source of inefficiency in an application system but was considered preferable in the framework the development of visual / spatial architectural representations for the computer and of the investigation of the applicability of related recognition techniques that are attempted in the dissertation.

The second way in which the application domain influences the kind of types recognized by CHQP is that the number of architectural plans that can be produced from a limited number of atomic part types and relationships between these parts is, for all practical purposes, infinite. Even if these part types and their relationships are quite few, they can still produce a vast number of plans if no additional acceptability constraints are introduced, as we can judge from a comparison between the Palladian grammar by Stiny & Mitchell [1978a, b] and the similar but underconstrained grammar induced by Mackenzie [1988].

The potentially infinite number of plan and group types suggests that an alternative to matching a description of the derived group configuration, such as the one in Figure 3-49, to stored models of parts or wholes of architectural plans, is evaluation of an architectural plan with respect to general formal constraints of acceptability. In other words, instead of integrating such constraints in the process of recognition, it was considered preferable to organize them into an independent module of CHQP, as this conforms to both the proscriptive manner of the classical canon [Tzonis & Lefaivre 1986, p.6] and to the way humans read architectural plans: ignorance of the formal constraints of the classical canon does not impede reading the articulation and subdivision of an architectural plan but simply reduces the accuracy and precision of reading and disallows comprehensive aesthetic appreciation of the observed patterns.



**Figure 3-49** A rudimentary relational graph that describes Villa Foscari in terms of its parts and their attributes

Matching of the final group configurations in Figures 3-44 to 3-48 to prototypical templates of location groups could have allowed for the recognition of instances of these types in the plans but could not directly determine the acceptability of these configurations with respect with what in different fields is referred to as *well-formedness* and *figural goodness*. Matching to models of whole plans would have probably resolved this problem but only implicitly, in the sense that there would be no explicit evaluation of the well-formedness or figural goodness of the group configuration, that is, of its conformity to the classical canon. And, since the models of the whole plans are essentially nothing more than a selective combination of different aspects of the classical canon, it was considered preferable to investigate the acceptability and overall articulation of group configurations derived by the process described in this chapter directly with respect to these aspects.

For these reasons the evaluation of descriptions of architectural plans and recognition of plan types in *CHAP* goes beyond matching and deserves a detailed discussion which occupies Chapter 4. A result of the distinction between the grouping process and the evaluation of the acceptability of its products is that, as the grouping process described in this chapter usually produces more than one alternative descriptions for the same architectural plan, it is not possible to offer at this point of the dissertation an evaluation of the effectiveness and reliability of the grouping process in terms of its correspondence of its output with

intuitive subdivisions of the plan and its spatial articulation. Therefore, this chapter concludes not with an evaluation of the process it describes but with a couple of remarks on specific problems encountered in the grouping process.

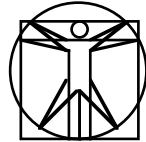
A major issue in Chapter 3 has been the extensive use of grouping relationships for the description and manipulation of location groups. On each level of grouping we have seen the complexity of the groups increase from simple aggregates of locations to complex groups characterized by relationships often other than these of their components and then to supergroups which may cover almost the whole plan. On some levels groups overlap, on others they may be fused or separated and their members connected to different groups. Any attempt to keep track of all these parts and subparts as explicit entities, of all attributes of parts and all local relations, is probably futile as the resulting description would become too cumbersome for any practical purpose. A shift from parts and groups to the relationships which characterize them allows a higher degree of abstraction, including attenuation of local perturbations and local changes of orientation and facilitation of group correlation. This emphasis on relationships relates to the perceptual tendency to resolve, for instance, projective ambiguities of stimuli by maximizing certain regularities. Perceptual sensitivity to each kind of regularity is characterized by preferences which relate to the preference order adopted in CHAP [Attneave 1982; Perkins 1982].

The abstraction provided by this emphasis on relationships in the descriptive structures of CHAP is helpful in many respects but cannot resolve by itself the primary computational problem of the grouping process described in this chapter, the lack of an effective control mechanism for the reduction of the number of possible combinations of groups on the second level of grouping. As a result, recognition of complex groups in the more complex architectural plans is not very efficient. Small plans, like those of the Palladian villas, have a small number of locations and hence a relatively small number of location groups on each level which makes them useful as examples because the reader can trace the recognition process in CHAP in a limited number of steps and pictures. However, compactness also results in many overlaps between groups and hence leads to more ambiguities on the second level of grouping than in larger plans. The development of a control mechanism not only for the second level of grouping but for the whole grouping process is seen as an issue related to the calibration of the procedures of CHAP: only a full implementation of the whole recognition and representation system proposed in the dissertation can establish the relative importance and priority of each selection or rejection constraint and rule at each stage of the grouping process and

thus indicate with accuracy the structure and composition of control mechanisms.

# CHAPTER 4

## COGNITIVE FILTERING



### 4.1 Introduction

#### 4.1.1 The interpretation of descriptions

Chapter 3 described the recognition of grouping relationships between locations, the atomic parts of architectural plans in *chcp*, which had been identified by the process described in Chapter 2. The descriptions finally returned by the process of Chapter 3 are relational structures of the parts of the architectural plan (i.e., mutually exclusive location groups). These structures are comparable to the ones returned by most computer vision systems [Brooks 1981; Marr 1982; Pentland 1986] and correspond (potentially at least) to the intuitive interpretations of grouping and subdivision in classical architectural plans, in very much the same way computer vision systems attempt to recover the configuration of the canonical or natural parts of objects and scenes depicted in an image.

In a computer vision system the obvious next step would have been to match the derived descriptions to general relational structures which represent types of whole architectural plans or types of plan parts [as in Marr 1982, chapter 5]. The latter would allow us to label and verify the recognized groups while the former would allow categorization of the described architectural plan in an explicit or implicit typology of prototypes, as in rectangular arrangements [Mitchell et al 1976; Steadman 1983], or of sequences of generative operations, as in shape grammars and syntactic pattern recognition [Gips 1975; Stiny 1975, 1976, 1980 — Duda & Hart 1973, chapter 12; Fu 1974, 1982, 1986; Gonzalez & Thomason 1978; Pavlidis 1977; Watanabe 1985, chapter 10].

Variations of this process of recognition by categorization constitute integral modules of all computer

vision systems. These modules allow the purposeful interpretation and manipulation of derived descriptions as it is recognition that assigns meaning to a description derived from an image. Recognition of parts and wholes provides the means for the coherent correlation of parts that is necessary for the completion of a description with missing explicit and implicit aspects and associations. For example, if a configuration of four slender cylinders and a rectangular slab on top of these is recognized as a table, then it is possible to infer that the four cylinders support the slab and are not suspended from it, or to make reasonable assumptions about the absolute size of each part and even about the materials it is made from.

From this viewpoint recognition by categorization amounts to the integration of extrinsic knowledge bases into the one used for the derivation and formulation of the descriptions. The extent and intricacies of this integration often pass unnoticed even though they have been the subject of influential studies in recognition [Winston 1975]. Perhaps their most direct expression can be found in the increase of required analysis of the model base proportionally to the increase of scope of a recognition system.

Character recognition offers a good example of this point. Character recognition of printed fonts has generally concentrated on global measurements for each character, ranging from the pixel templates that are currently employed in optical character recognition to the early magnetic sensor measurements used in conjunction with specially designed fonts. These measurements rely on the standardized appearance of a character within each font. In multi-font systems skeletonization and other abstraction techniques reduce the variability that results from the differences between different fonts and usually allow for the use of a single prototypical or reference pattern for each character. When, however, we cross over to handwritten text we see that abstraction and global measurements do not suffice in the face of the great variability and ambiguity that characterizes the appearance of each character. A popular approach to the resolution of this problem has been the decomposition of both the characters under investigation and of the stored prototypes into relational structures of strokes, joints and other parts. This allows greater flexibility in recognition and offers direct or indirect connections with constraints that determine the changes in direction and the temporal order of each part of a character in handwriting.

To return to *chap*, the example of character recognition suggests that a fruitful approach to confronting variability is to analyse further both derived descriptions and reference patterns so as to make explicit the underlying structure of each description and pattern and thus associate it with constraints that are not directly involved in the

perceptual organization and description of an image. Such analysis can be particularly beneficial in cases where the number of patterns that may be encountered is very large. Architectural plans, even when considered within a single formal system, are a typical case: although at a suitably high level of abstraction one may suggest a potentially finite set of plan type categories, there is no way to achieve a complete enumeration of all types that fall within each category because even a quite limited number of elements and relations may produce an infinite number of combinations. The only safe predefined typologies are derived inductively from a limited number of instances, as in the grammar of Palladian villas [Stiny & Mitchell 1978a, b]. If, however, we attempt to cover, for instance, the whole of classical architecture or even just the whole production of Andrea Palladio we need to refer to and take account of the more general and abstract principles of classical architecture which do not constitute deterministic or even well-defined systems [Summerson 1980; Tzonis & Lefaivre 1986].

An alternative approach is to attempt to arrive at a taxonomy of architectural plans not by matching to particular instances or types but by analysing the derived descriptions through general testing devices towards the identification and evaluation of certain principles and their contribution to the structure of the plan. This approach allows direct reference to the aforementioned general categories of types or rather to their constituent constraints, which can also be obtained through similar analysis. Such *analytic* investigation of general constraints rather than patterns is the subject of this chapter. The adoption of this approach means that in CHQP meaning is not attached to descriptions of architectural plans by reference to general types which are superimposed to the descriptions of particular instances. Instead, *interpretation of descriptions of architectural plans relies on the analysis of these descriptions with respect to the principles which qualify an architectural plan as a classical composition.*

Preference for such an analytic approach over matching does not dispute the power of interpretation that is implicit in direct classification. It merely implies that a successful classification system integrates in fact several analyses which provide it with power of interpretation. In one celebrated case of successful classification, Mendeleev partially modified the order of elements by their atomic weight so as to take into account their chemical properties. The fact that the resulting periodic table is based on the atomic number of elements should therefore be considered as a concise expression of the underlying principles that consistently explain the results of the analyses involved in the definition of the table and not as the basic taxonomic criterion.

The adoption of an analytic approach also relates to studies on human perception which suggest the existence of a number of alternative relational configuration of parts, multiple reference frames and multiple rule structures which interpret a scene often regardless of the tasks in hand. All these appear to be instruments of an extremely complex strategy whose purpose is to compensate for inaccuracies, errors and lapses [Perkins 1983] and thus lead to perceptions which can be explained by a common cause [Rock 1983].

The adoption of this analytic approach at this stage of *chap* aims at something equivalent. Instead of superimposing a self-contained classification model to the descriptive process of Chapter 3, the extent and significance of the primary classification criteria is investigated in the resulting descriptions of architectural plans. This elaboration of the significance and the applicability of each criterion aims (a) at the development of an accurately and precisely calibrated system of typologic criteria for architectural plans, and (b) at the formulation of a bottom-up classification process that is closely correlated with the structure of the bottom-up derivation of the description and hence allows for a smoother transition from derivation to matching to known models.

#### 4.1.2 An outline of the chapter

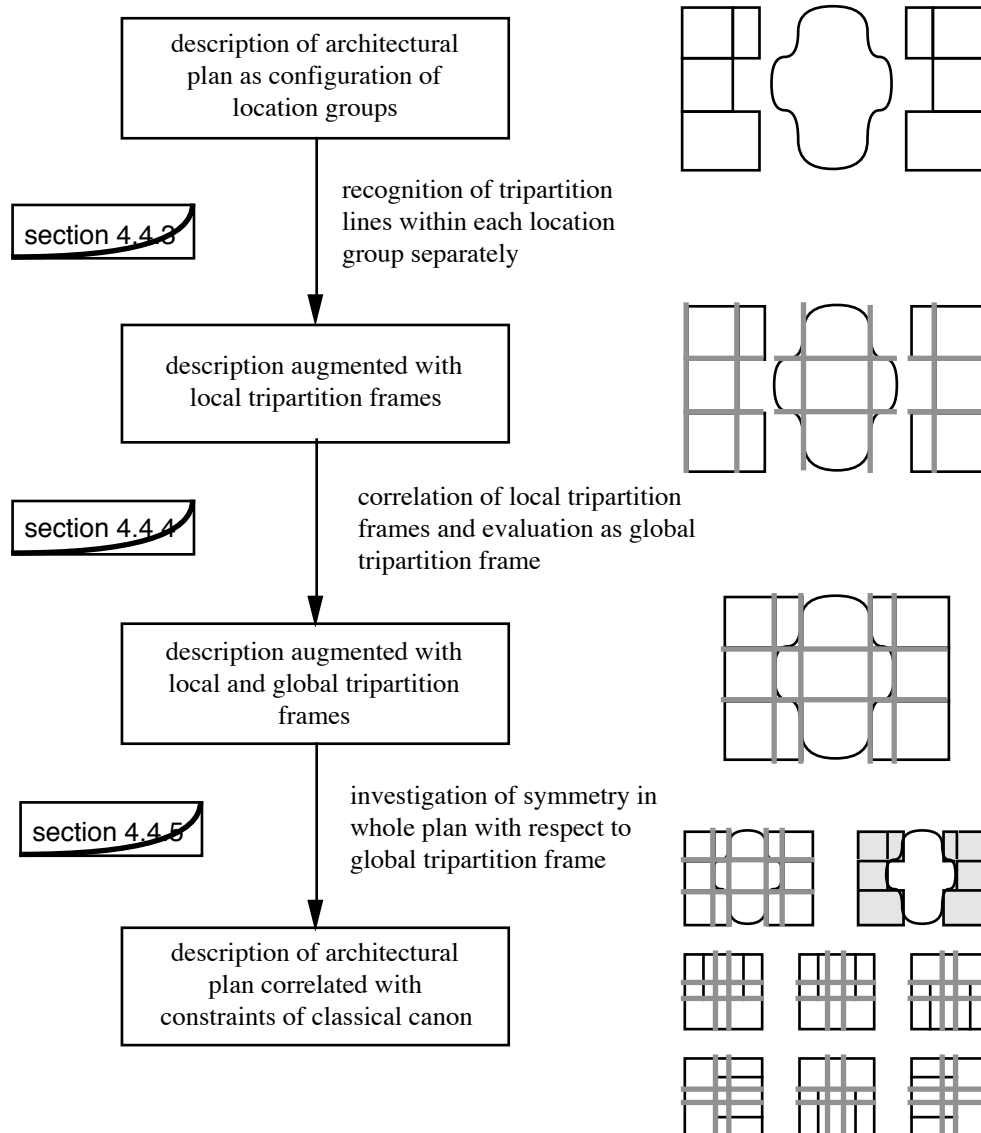
The present chapter represents a departure from the ‘lexical’ level of the description and recognition of architectural plans we have considered in the previous chapters towards a ‘syntactic’ level where classification and evaluation of alternative descriptions of the same plan as well as of descriptions of different plans are the main concerns.

Analysis of the alternative descriptions of an architectural plan with the purpose of establishing a preference order for these alternatives is related to the issues of figural goodness and well-formedness. Section 4.2 considers two alternative approaches with respect to these issues: (a) Gestalt-related and information theoretic measures of figural goodness, and (b) comparison to the general principles which qualify a building as a classical composition.

In *chap* the analysis of the alternative descriptions of an architectural plan with respect to well-formedness follows the second approach and utilizes sequences of *cognitive filters* which analyse the particular expressions of the constraints which comprise the canon of classical architecture with the purpose of distinguishing between well-formed instances and random or accidental agglomerations

both on the level of alternative descriptions and on that of the actual architectural plan.

The proposed technique of cognitive filtering is introduced in section 4.3 through an extension of Marr's [1982] computer vision approach which facilitates recognition of the posture of non-rigid forms. The purpose of this extension is to show that decomposition of the model base offers advantages even in cases where recognition implies very strongly matching to complete models. Recognition of posture is also related to cognitive filtering in chap in another sense: as recognition of posture represents an elaboration of the recognition of a model such as a human being or an ape, investigation of the conformity of the description of an architectural plan goes beyond recognition of an architectural plan or of a particular plan type and therefore provides explicitly many of the aspects that should concern manipulation of architectural plans in an intelligent computer system.



**Figure 4-1** A schematic outline of the module of chap described in Chapter 4 with respect to the general outline of Figure 1-1

Section 4.4 discusses the cognitive filtering system used in chap and its relations with the formalization of the classical canon by Tzonis & Lefaivre [1986]. While the generative system suggested by Tzonis & Lefaivre is top-down, cognitive filtering in chap proceeds in a bottom-up manner. This difference allows for a more comprehensive verification of this formalization as it considers the classical canon from the viewpoint of perception and recognition instead of the (more usual in CAAD) perspective of problem solving. The structure of cognitive filters in chap follows the analysis of the classical canon into schemata (taxis, genera and symmetry) and subschemata and investigates those relevant to the scaled-down context of chap on a local and a global level, that is, within each location group of the description and for the whole plan.

Section 4.5 effectively concludes the presentation of chap by presenting potential extensions of the present

version with respect to the structure and purpose of its most unconventional part, cognitive filtering. Cognitive filtering is proposed as a device for accommodating the great variety of points of view and concerns involved in the selection of a particular spatial arrangement in architectural design and as an approach to the integration of matching to known models into the derivation of the description of an image.

Figure 4-1 presents a schematic outline of the process described in Chapter 4 with respect to the outline of chap in Figure 1-1.

## 4.2 Criteria of well-formedness

### 4.2.1 The evaluation of figural goodness

Cognitive filtering, the technique introduced in this chapter, essentially amounts (within chap) to an investigation of the acceptability of an architectural plan with respect to the canon of classical architecture. Therefore, cognitive filtering relates to investigations of figural goodness, that is, of the preference for certain patterns on the basis of their structural characteristics. These investigations generally derive jointly from the Gestalt notion of *Prägnanz* and from the notion of redundancy in information theory: by viewing the perceptual system as an information processing channel of limited capacity we may assume that more redundant patterns, such as symmetrical shapes, can be more efficiently and economically encoded and thus also processed and stored in memory than less redundant ones [Palmer 1982, p.97].

The notion of *Prägnanz* in Gestalt theory determines the best possible organization of stimuli under the prevailing conditions. In an analogue model *Prägnanz* corresponds to a tendency towards a goal state or a state of equilibrium and in a symbolic model to the preference for a simple description of the stimuli [Beck 1982a, p.2]. Under the influence of information theory symbolic models have attracted more attention, despite the fundamental problem that search for the most economical encoding requires that all permissible interpretations of the stimuli have been clearly recognized and somehow encoded. Given the very high number of such alternatives in perception, this seems implausible in human perception [Attneave 1982; Rock 1983, p.163] and futile in computer vision.

The alternatives to exhaustive enumeration of permissible interpretations include representation systems that seek directly the most economical description, such as hill climbing systems equipped with feedback mechanisms by

which the complexity of a description drives the system [Attneave 1972] or neural dipoles interacting towards conformity through mutual facilitation of like values or variables and inhibition of unlike ones [Attneave 1982]. Another approach is the use of preprocessing to reduce the number of permissible interpretations, for instance by linking tokens of the basis of proximity and connectedness [Beck 1982b] or by preferentially maximizing certain geometric constraints [Perkins 1982].

In this context, the grouping module of CHQP (cfr. Chapter 3) can be considered as a case of the latter approach, that is, as an organization of the image into a limited number of alternative descriptions on the basis of certain general constraints. This allows us to ignore the potential multiplicity of possible interpretations of the same pattern and appreciate formally each resulting description independently, that is, investigate the well-formedness of each description and thus the well-formedness of the architectural plan it describes.

Quantitative measurement of figural goodness was initiated by Attneave [1954] and Hochberg & McAlister [1954] who proposed independently that from the viewpoint of information theory good figures contained less information (i.e., were more redundant) than bad ones. Of the antecedents of this line of thought we shall be concerned with two quite distinct approaches: *transformational economy* [Garner & Clement 1963; Garner 1974; Palmer 1982, 1983] and Leeuwenberg's [1971, 1978] *coding theory*. The former will be discussed in subsection 4.4.5, in the framework of the analysis of symmetry in architectural plans. The applicability of the latter to this stage of CHQP will concern us in the present subsection because, as an advanced and well-defined system for the measurement of information in a given pattern which overcomes the difficulties of earlier information theoretic approaches, it provides a proper measure of the potential of measuring the well-formedness of architectural plans in terms of information theoretic economy.

In architecture the interest in expressions of figural goodness has been rather limited, presumably due to the generally functionalist overtones of research in architectural problem solving. Despite the influences of Gestalt theory and even more of information theory on many studies on formal analysis and composition, there have been few attempts to evaluate aesthetics of buildings on the basis of rigorous Gestaltist or information theoretic criteria [Prak 1977].

Perhaps the closest to a consistent evaluation system for a fair proportion of the qualitative formal constraints of a particular design approach can be found in shape grammars [Gips 1975, section 3; Stiny 1975, part II; Stiny & Gips 1978a] and in particular in the evaluation of Palladian plans [Stiny & Gips 1978b], that is, of plans produced by the

Palladian grammar [Stiny & Mitchell 1978a, b]. In this, Palladian plans are evaluated formally on the basis of (a) ad hoc local criteria inductively derived from the plans of Palladio's actual villas and (b) a global aesthetic measure. The former are admittedly indistinguishable from the generative shape rules of the Palladian grammar and concern either individual locations or groups of cells of the underlying 3 x 3, 5 x 3 or 5 x 4 grid. The global aesthetic measure is defined as the ratio of the length of the description of a plan (i.e., the sum of the number of cells required for the multicell location types of the plan and of the number of instances of each of these types in the plan) to the length of the information required for its generation (i.e., the number of shape rules required for the generation of the plan).

The local criteria, being implicit expressions of formal and functional constraints of Palladian composition, are more or less effective with respect to the ultimate criterion of well-formedness one may expect for Palladian plans, that is, the distinction between the actual plans by Palladio (plus some other ones) from similar ones produced by the Palladian grammar. The global measure, on the other hand, is "a measure of the *specificational simplicity* of a plan" [Stiny & Gips 1978b, p.204] which therefore says more about the particular formative approach implied by the Palladian grammar than about the formal quality of the plan. In other words, although the system is quite successful in the evaluation of the products of the Palladian grammar, it loses almost any meaning and purpose outside shape grammars because a shape grammar can account for a single formative history and only implicitly for the general constraints of classical architecture which form the general framework of Palladian composition.

A more general formal evaluation system is provided by Leeuwenberg's [1971, 1978, 1982; Buffart et al 1981] *coding theory*. In coding theory a pattern is first described in terms of an alphabet of a number of atomic primitive types, such as straight line segments and angles at which the segments meet. This description (the *primitive code*) carries an amount of structural information that is equal to the number of elements (i.e., instances of the primitives) it contains. The structural information of the primitive code is then minimized by repeatedly transforming the primitive code on the basis of a limited number of coding operations (iteration, reversal, distribution and continuation) until the *end code*, a code whose structural information cannot be further reduced, is reached. The structural information of a pattern is that of its end code.

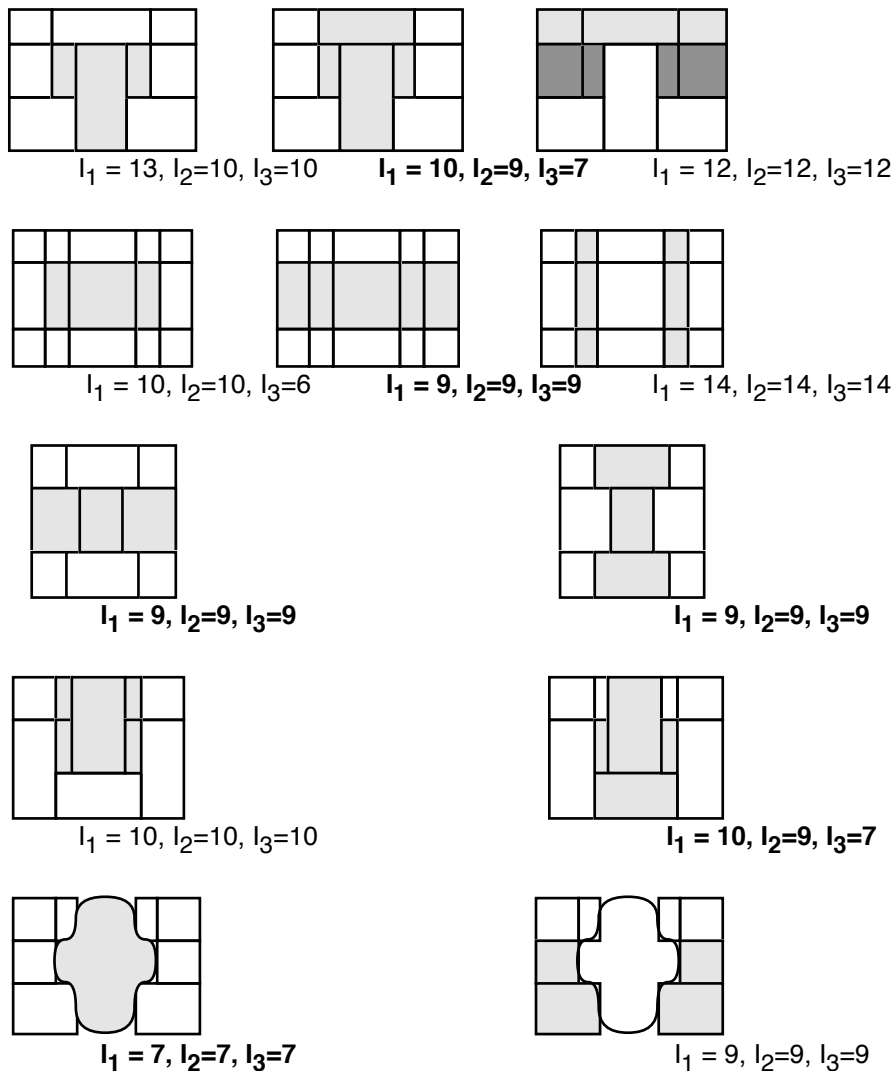
A proper application of coding theory to architectural plans goes beyond the scope of *CHQP*. Within the limitations of the present investigation of figural goodness a

rough version is attempted on the basis of an alphabet of (a) the location shape types that may be encountered in rectangular Palladian villa plans (cfr. subsection 3.2.1) and (b) the contiguity relationships between locations as described by the adapted chain coding scheme (Figure 3-6).

The structural information of each of the final group configurations of the examples in Chapter 3 was measured as the sum of the structural information of the groups in the configuration (Figure 4-2). For this, three variations of the coding alphabet were used. In the first ( $I_1$  in Figure 4-2), each of the twelve distinct contiguity relationship of adapted chain coding is described by a distinct primitive and pairs of groups that are completely symmetric with respect to a central horizontal or vertical axis are treated as a single iterative group. This means that in calculating the structural information of the whole configuration, instead of multiplying the structural information of the one group by two, a single bit is added to the structural information of the one group. In the second variation ( $I_2$  in Figure 4-2) the number of contiguity relationships of adapted chain coding are reduced to eight by considering pairs of relationships that are symmetric with respect to a horizontal or vertical axis to be subtypes of the same type. For example, relationships 11/13 and 12/14 are assigned to the same primitive of the coding alphabet. The result of this modification is a non-uniform reduction of the structural information of each group configuration. The third variation ( $I_3$  in Figure 4-2) reduces further the number of contiguity relationships to just four by ignoring the direction of articulation. For example, relationships 11/13, 12/14, 21/23 and 22/24 are assigned to the same primitive of the coding alphabet.

Although a proper evaluation of the three variations and of the applicability of coding theory to CHQP cannot be based on just these examples, it is evident that the measurement of structural information of location groups that comprise a description avoids the limitations imposed by generative systems because it operates on a declarative level and hence can be superimposed to any descriptive system. A proper adaptation of coding theory could measure certain major formal properties of architectural plans, such as its compactness and the homogeneity of its parts but would still provide only implicit views of the classical canon, that is, it would not measure explicitly an architectural plan against the constraints of its general formative framework. It is highly probable that coding theory can be used to select the best among a number of alternative descriptions of an architectural plan, in the same way it has been used to explain quite a few of our perceptual preferences. Its success in architecture would, however, rely entirely on the alphabet used, as this alphabet

would constitute an indirect expression of the general constraints of classical architecture. Such an alphabet presupposes rigorous analysis and quantification of the constraints of classical architecture which could also lead to the development of a methodology of direct formal evaluation of classical plans, albeit in a qualitative manner.



**Figure 4-2** Evaluation of alternative descriptions of architectural plans (cfr. Figures 3-44 to 3-48) by a rough application of coding theory

$l_1$  denotes the first version,  $l_2$  the second and  $l_3$  the third version of this application; bold typeface denotes the preferred description (on the basis of the second version)

## 4.2.2 The classical canon

An alternative way of investigating the well-formedness of an architectural plan is in direct comparison to the abstract principles and norms of its aesthetic framework, which for the initial version of CHAP is classical architecture. This is not a novel approach, as, in one version or another, it has

been the dominant technique in architectural history and theory. As Summerson [1977, p.8] has put it, classical buildings have been traditionally evaluated aesthetically either by their conformity to a set of given classical orders or by their correspondence to some abstract 'essence' of these orders as a whole or by a combination of both. Such intuitive analysis or analysis on the level of building elements does not apply to CHQP. What is required is a set of explicit, computable rules which offer a measure of the well-formedness of classical architectural plans on the basis of the descriptions returned by the process discussed in Chapter 3.

Generative systems like shape grammars generally avoid confronting directly the issue of well-formedness with respect to some high, abstract principles. Instead, consistent with their generative nature, they integrate facets of such principles into the various (local) formative operations or into local evaluation criteria, as in the evaluation of Palladian plans by Stiny & Gips [1978b]. It is indisputable that a properly calibrated generative system, such as the Palladian grammar, can be quite successful in producing architectural plans that are consistent with a specific set of principles, such as the underlying formal principles of the villas designed by Andrea Palladio. However, the consistency of such systems is undermined by the lack of a general coordinating formal structure which, although implicit in an architectural plan or an actual building, guides our perceptions in a powerful and sometimes even deterministic manner.

Investigation of the well-formedness of an architectural plan in CHQP adopts an opposing point of view, namely that local operations and conditions are dependent upon general formal coordinating devices. The two basic purposes of the investigation of well-formedness in CHQP are (a) the identification of a coherent general organizational structure that relates an architectural plan to the classical canon, and (b) through this, the organization of the configuration of location groups returned by the descriptive process, in the same way that deictic reference frames help organize our cognitive maps.

A prerequisite to such an investigation is the adoption and adaptation of a domain theory, that is, a formalization of the classical canon. In its treatment of domain theories and domain knowledge in general the dissertation is characterized by what we could name the knowledge engineering perspective: CHQP accepts a specific formalization of the classical canon as a coherent body of domain knowledge for reasons that have more to do with the structure of CHQP and its underlying approach to recognition than with architectural theory. Choice of a model of the classical canon relies on the computability, the potential of multilevel abstraction and the compatibility of

the model to the bottom-up recognition process of CHQP. An appropriate model should therefore (a) be computationally oriented, (b) provide the structure and composition of formal coordinating structures on a variety of interpretation and abstraction levels, and (c) offer the means to distinguish between the description or generation of an architectural plan and its evaluation with respect to the classical canon. CHQP has been developed specifically with reference to a formalization of the classical canon that meets these specifications, that by Tzonis & Lefaivre [1986]. Acceptance of this formalization from a knowledge engineering perspective means that CHQP is not bound by it as any compatible domain theory is equally acceptable, as any additional or alternative body of domain knowledge can be, in principle at least, integrated into a knowledge based system.

The compatibility of different accounts of the classical canon for CHQP depends to a large degree on their notion of coordinating formal structures in a classical design. Attempts to make explicit coordinating formal structures are not unaccustomed in studies of classical architecture. Wittkower [1952, Part III, Figure 8] describes the plans of Palladio's villas with reference to the 5 x 3 grid pattern which he calls the "geometrical patterns of Palladio's villas" [p.65], the pattern which epitomizes Palladio's rules of geometric arrangement. He suggests that all villa plans "are derived from a single geometrical formula" [p.64] but points out that "this grouping and re-grouping of the same pattern is not as simple an operation as it may appear" and, in a manner characteristic of the general resistance to accept the 5 x 3 grid pattern as a primeval generic form from which all villa plans are directly derived, he stresses the role of "harmonic ratios not only inside each single room, but also in the relation of the rooms to each other" [p.66]. Another point to which we shall return later is that Wittkower relates this pattern to "the straightforward needs of the Italian villa" [p.64] in enclosed and semi-enclosed spaces.

The same line of thought appears to underlie the Palladian grammar [Stiny & Mitchell 1978a, b] where the composition of a villa plan takes place in three stages: grid definition, exterior wall definition and room layout. Grid definition invariably results into a 3 x 3 or 5 x 3 grid pattern, yet these patterns are not taken for granted. Instead, they are produced by sequences of shape rules which concern the relative position of pairs or triads of locations. Exterior wall definition simply amounts to circumscribing the grid by a rectangle to form the wall pattern that corresponds to the 3 x 3 or 5 x 3 grid, i.e., the underlying wall patterns for these two classes of floor plans. The members of the classes are each derived in the third stage, room layout. In this stage, grid cells are recursively fused to form the locations of the plan.

All these locations have rectangular shape, except for the central one which may have the shape of a cross, a T or a double T (H rotated by 90°). In this framework the 3 x 3 and 5 x 3 patterns are taken not as coordinating structures but rather as parts of the raw materials from which each distinct plan is formed.

Objections to the exhaustive enumeration implied by shape grammars and to the local character of their rules come from a historical perspective: Ackerman [1977, p.36] observes that among Palladio's villas "there are few instances of a repeated plan ...; Palladio would produce at most two or three versions of a particular scheme before reaching out in an entirely new direction. The common core within this variety is a particular conception of architectural harmony and composition." Therefore, Ackerman concludes, there is no typical Palladian villa. There is, however, a uniform schema which underlies plans and elevations: "a triadic composition with a central block built around the axis of the entranceway, and two symmetrical flanking blocks" [p.160]. This triadic system is according to Ackerman a coordinating structure which facilitates the unified use of proportionalities and also conforms to the traditional centralized arrangement of the typical Venetian house as elaborated by Palladio — what Wittkower [1952, p.64] called "the straightforward needs of the Italian villa".

- 1 taxis (framework) \*
  - 1.1 grid \*
    - 1.1.1 rectangular \*
    - 1.1.2 polar \*
  - 1.2 tripartition \*
    - 1.2.1 addition °
    - 1.2.2 deletion °
    - 1.2.3 fusion °
    - 1.2.4 embedding °
- 2 genera (elements) °
  - 2.1 Doric ·
  - 2.2 Ionic ·
  - 2.3 Corinthian ·
    - 2.3.1 column ·
      - entablature ·
      - cornice ·
      - frieze ·
      - architrave ·
      - column ·
        - capital ·
        - shaft ·
        - base ·
      - stylobate ·
      - pedestal ·
    - 2.3.2 pillar ·
    - 2.3.3 pilaster ·
    - 2.3.4 window ·
    - 2.3.5 door ·
    - 2.3.... . . . ·
  - 2.4 Tuscan ·
  - 2.5 composite ·
- 3 symmetry (relations) \*
  - 3.1 rhythm \*
    - 3.1.1 stress °
    - 3.1.2 contrast °

- 3.1.3 reiteration °
- 3.1.4 grouping °
- 3.2 figures \*
- 3.2.1 overt \*
  - parallelism °
  - analogy °
  - alignment °
  - contrast °
- 3.2.2 subtle \*
  - aposiopesis °
  - abruptio °
  - epistrophe °
  - oxymoron °
  - ... °

**Figure 4-3** *The classical canon according to Tzonis & Lefaivre [1986]*

(\* : pattern [abstract regulating schema or primitive configuration];  
 ° : procedure [for the derivation of patterns from generic entities];  
 · : element [standardized — even if only parametrically—  
 configuration])

The concept of a coordinating structure is further investigated by Tzonis & Lefaivre [1986] in their study of classical architecture as a formal system. This study provides the essential domain knowledge utilized in CHAP for the evaluation of well-formedness and will therefore be presented in detail. The reasons for the adoption of this particular formulation of the classical canon is that it is analytic, precise, computationally oriented and proscriptive rather than prescriptive: Tzonis & Lefaivre suggest that the constituents of the canon “do not so much direct action as *constrain* it ... . In other words, instead of telling us what to do, they tell us what not to do. This may seem to be a subtle equivocation. But it explains why so many new classical formal arrangements have been, are, and probably will continue to be created out of the same canon. By constraining rather than directing, the classical canon allows for a certain degree of freedom and invention in responding to those forces of change that lie outside the world of forms” [p.6].

This proscriptive character has important consequences for CHAP because it justifies the explicit consideration of classical constraints after grouping of locations has been completed, i.e., the consideration of the classical canon not as a deterministic generative system but as a controlling and coordinating formal structure that determines the formal quality of an architectural work and not its feasibility as a container of certain activities. This is reinforced by the explicitness of this formal structure and its detachment from local grouping or other formal conditions. A final merit of the formulation suggested by Tzonis & Lefaivre is that, in contrast to the mainstream of the current re-kindling of interest in architectural forms of the past, it goes beyond the appearance of buildings and concentrates on fundamental issues of spatial arrangement and organization, that is,

precisely on factors that determine our perceptual and operational appreciation of a building. In other words, classicism or any other architectural formal system is a matter of composition rather than decoration. A building with an ionic portico is not necessarily classical, while the absence of classical detailing does not disqualify completely a building from being classical (Summerson's [1980, p.8] objections withstanding).

Tzonis & Lefaivre [1986] suggest that the classical canon consists of three major levels, *taxis*, *genera* and *symmetry*. *Taxis* is responsible for the overall organization of a building: "Taxis divides a building into parts and fits into the resulting partitions the architectural elements, producing a coherent work" [p.9]. *Genera* (the term preferred over the traditional "classical orders") are the "well-determined sets" of architectural elements which are formed on the basis of "particular fixed relations" [p.35]. *Symmetry* should not be confused with the mathematical notion of perfect correspondence between elements with respect to an anchor entity. Tzonis & Lefaivre use the term in a broader sense that also means more than the Vitruvian denotation of commensurability of elements: "Here, symmetry is used to cover universally all constraints of architectural composition that refer to how elements are chosen and placed in relation both to one another and to the overall structure of taxis" [p.117]. Cognitive filtering in *CHAP* concerns *taxis* and *symmetry*, since all references to *genera* are eliminated by the fact that our descriptions and analyses concern abstracted binary images of architectural plans.

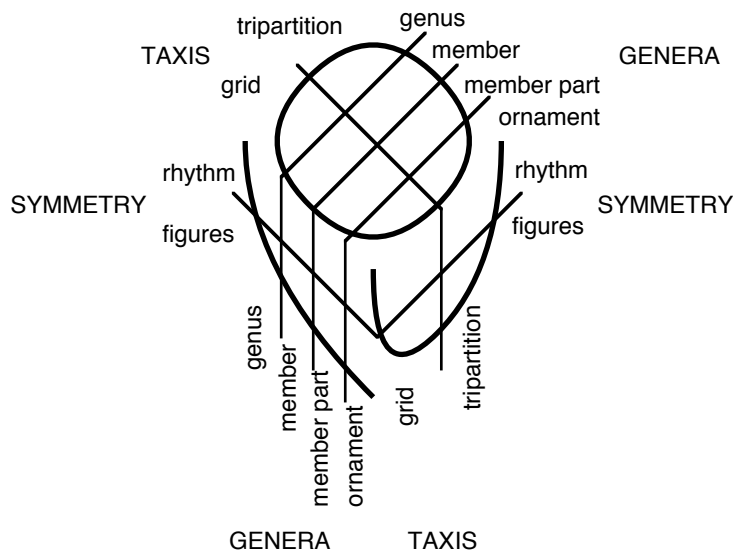
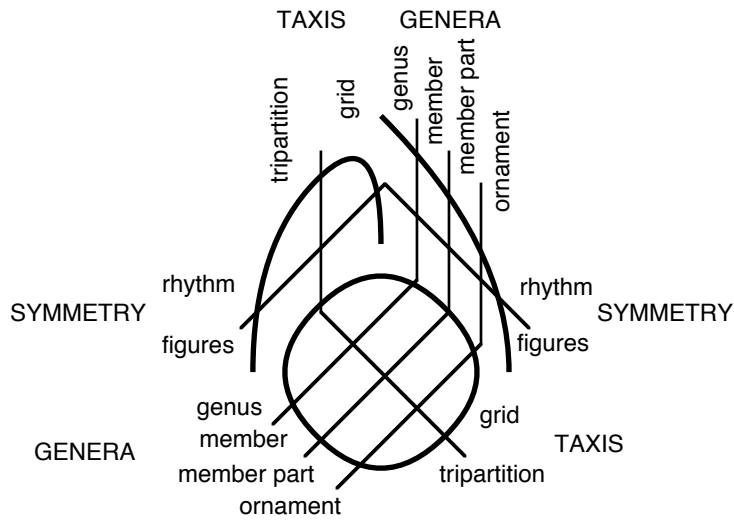
*Taxis* "contains two sublevels, which we will call *schemata*: the *grid* and the *tripartition*. The grid schema divides the building through two sets of lines" [p.9]. It can be described roughly as a parametric coordinate structure which regulates the positioning of building elements and thus the physical appearance of its more general and fundamental partner, *tripartition*. "The schema of *tripartition* marks the difference between the internal and external sections of the work. It divides a building into three parts and one enclosed" [p.14]. The notion of the grid is not used extensively in *CHAP* because the current version deals with totally orthogonal plans only and hence contains a strong though implicit rectangular grid. Moreover, the *tripartition* schema appears to be embodied in the generic 'square and cross' or nine squares pattern,

a	b	a
b	c	b
a	b	a

i.e., the 3 x 3 grid pattern, which in the above form also contains an implicit rectangular grid. From this pattern abstract plan layouts are derived in a



This supremacy of taxis becomes evident if we attempt to transform this descriptive analysis of the classical canon (Figure 4-3) into a correlation of the three levels which reflects the relative influence of each pair of levels (Figure 4-4). Consequently, taxis, and in particular tripartition, becomes the obvious starting point for the investigation of well-formedness in chq̄p.



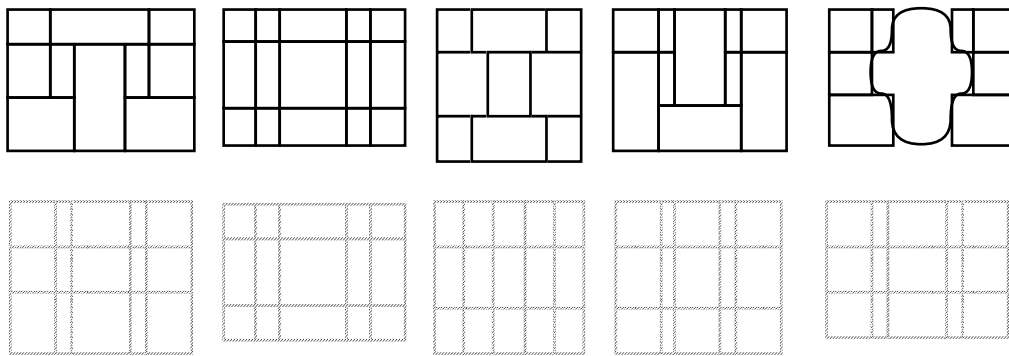
<b>SYMMETRY</b>	<b>rhythm</b>	modular coordination	pattern organization (grouping & subdivision)		
	<b>figures</b>	spatial allocation	correlation of figure components		
<b>GENERA</b>	<b>genus</b>			correlation of components	repetition (continuation) of components
	<b>member</b>	modular coordination	subdivision	correlation of components	repetition (continuation) of components
	<b>member part</b>	modular coordination	subdivision	correlation of components	repetition (continuation) of components
	<b>ornament</b>	modular coordination	articulation	correlation of components	repetition (continuation) of components
		<b>grid</b>	<b>tripartition</b>	<b>rhythm</b>	<b>figures</b>

**Figure 4-4** A correlation of the levels of the classical canon (cfr. Figure 4-3)

The transformability of the 3 x 3 pattern into every possible classical architectural plan through the recursive application of a small number of transformation rules lends itself to the suggestion that it would be possible to develop a predefined typology of plan types in a manner similar to that of shape grammars or rectangular arrangements. Through this typology a model base similar to that of most computer recognition systems could be developed and the models matched directly to descriptions of architectural plans, as in conventional pattern recognition systems [Ettinger 1988]. This approach has been repeatedly dismissed so far in chap and is dismissed once again because of the inefficiency that would result from the size and complexity of the model base

and also because it does not correspond to the analytic way we read plans.

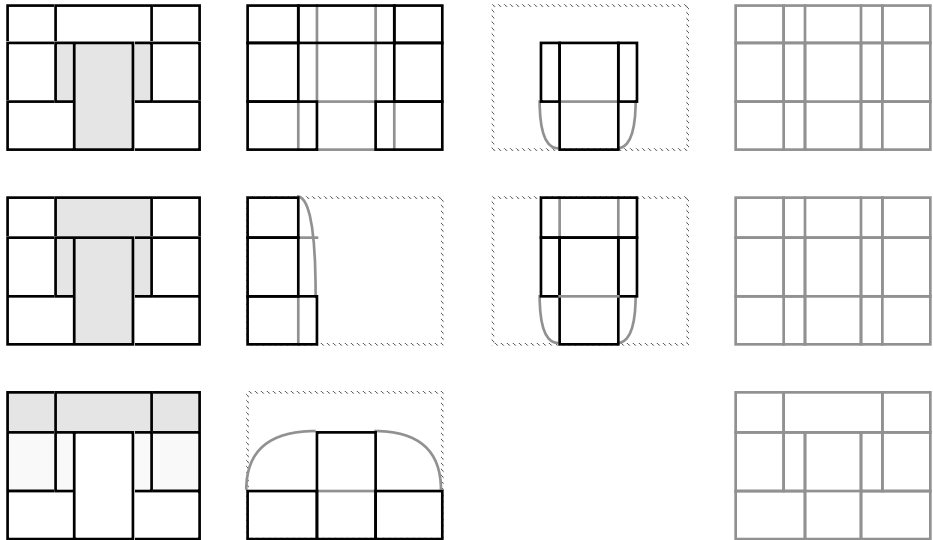
An alternative approach is to attempt to recover the 3 x 3 pattern or one of its derivatives from the plan. As this pattern or one of its derivatives forms an underlying grid to every classical architectural plan, a sensible way to make it explicit is through the use of *gratings* [March 1976]. By extending the edges of each location until they meet the perimeter in the architectural plans we have used as examples in Chapter 3 we invariably obtain a 5 x 3 grid (Figure 4-5).



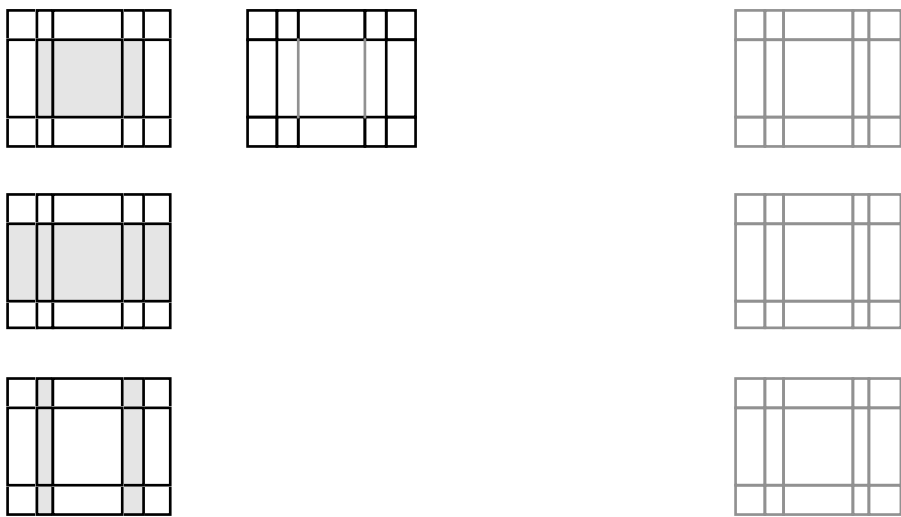
**Figure 4-5** *Minimal gratings of whole architectural plans*

This global investigation obviously tells very little about the acceptability of an architectural plan as a classical composition beyond that it is based on an acceptable grid and even less about the spatial arrangement of the plan or about alternative descriptions of the plan, such as the final group configurations derived in Chapter 3 (Figures 3-44 to 3-48). One solution is to obtain the minimal grating of each location group of each description and derive the grating of each description by collating the gratings of its groups (Figure 4-6).

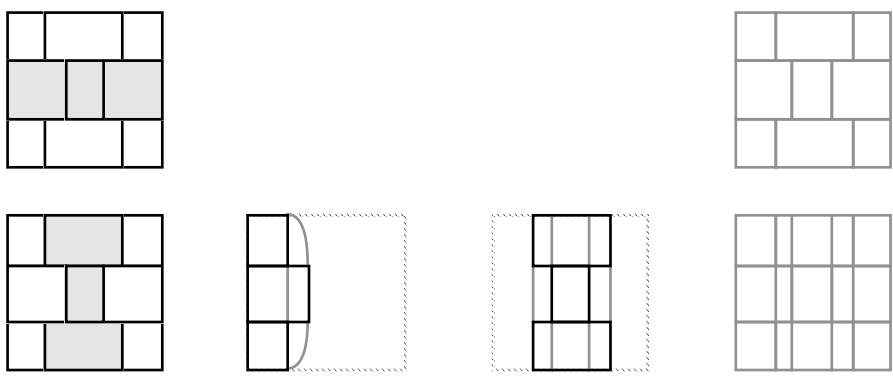
Villa Zeno



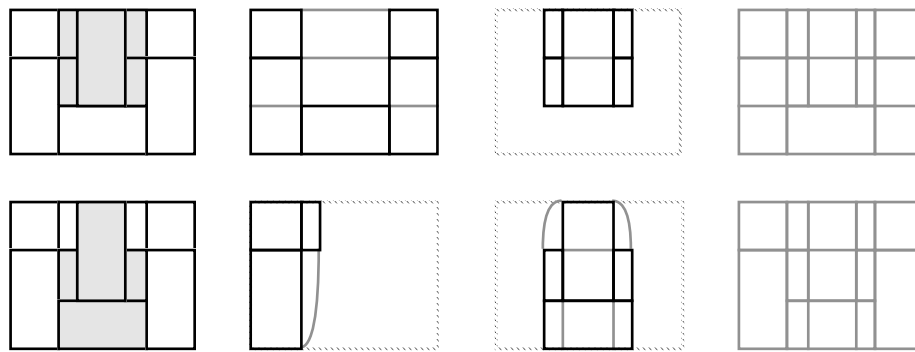
Villa Thiene



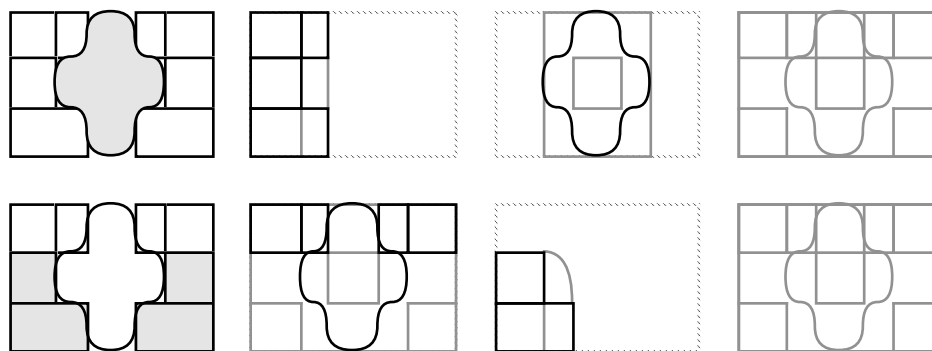
Villa Ragona



Villa Poiana



Villa Foscari



**Figure 4-6** Gratings of architectural plans produced by collating the minimal gratings of their groups (cfr. Figures 3-44 to 3-48)

This technique provides an indication of the articulation of each group and of each alternative description. As the grating of the whole plan is not necessarily a 3 x 3 or 5 x 3 pattern, it seems possible to devise a typology of classical plans on the basis of these gratings. By assigning some form of value to the grating of each part and to that of the whole plan, we would arrive at an evaluation of the well-formedness of classical plans that would allow (a) the consistent selection of a preferred plan description from its alternatives, and (b) the measurement of the well-formedness of an architectural plan. However, these possibilities are undermined by the dependence of gratings on the size and complexity of location groups. As one can see in Figure 4-6, the largest possible groups in Palladian villas (i.e., groups that contain all but three of the locations of the plan) generally have a 5 x 3 grating which usually corresponds to the grating of the whole plan. This grating suggests that we would have to accept these groups as being well-formed and put the descriptions they belong to in a high preference order. However, it is obvious that the 5 x 3 grating, is in these cases at least, not so much an indication of the well-formedness of the group as of its size and of the number of directions of articulation that characterize it. Reducing the permissible maximal size of a location group would not be a solution as it would disqualify certain group types of undeniable unity, such as the uniform loop of locations that often surrounds an atrium.

As a result, it is difficult to relate the resulting grating of an architectural plan with a particular plan type (i.e., model) or to use characteristics of such gratings in order to determine preference for one of the alternative descriptions of an architectural plan. In fact, the structural information measures of Figure 4-2 appear to offer more consistent, reliable and acceptable results even though they cannot be directly justified on the basis of formal aspects of the classical canon.

We are therefore confronted with a problem of the greatest magnitude in recognition that can be succinctly stated as how to relate and combine top-down model bases, such as the typology of architectural plans that can be inferred from Tzonis & Lefaivre, with essentially bottom-up descriptive processes, such as the one described in the previous chapters. Solutions proposed in Artificial Intelligence and related fields concentrate on sophisticated techniques which provide appropriate structure to both the derived description and the models [Minsky 1975] or elaborate and improve matching of the derived description to the stored models [Winston 1975]. These solutions increase the efficiency of the connection but do not result into an integration of the model base and the descriptive process.

Characteristic examples of the problems associated with this lack of integration can be found in probably the most comprehensive approach to computer vision to date, that of Marr [1982]. An extension of this approach towards the integration of the model base and thus of recognition in the descriptive process is presented in the following section (4.3). This extension introduces the technique employed at this stage of CHOP, *cognitive filtering*. Cognitive filtering is in a sense a reversal of the traditional approach to matching descriptions to models: instead to considering the models as integral entities and attempting to bring the descriptions to a state and form comparable to those of the model, models are decomposed into hierarchies of filters which operate on a local or global level. These hierarchies are superimposed to the descriptive process and, being structurally equivalent to derived descriptions, allow a progressive recognition of the model that corresponds to the description. This type of integration has the primary advantage that it allows greater flexibility to both the model base and the recognition process without destroying the coherence and the possible combinatorial character of the domain knowledge the model base derives from, as is the case with the formalization of the classical canon by Tzonis & Lefaivre.

## 4.3 Cognitive filtering and recognition of posture

### 4.3.1 Recognition within Marr's framework

Understanding of the structure and purpose of cognitive filtering in CHQP can be improved by considering first an analogous problem that is more straightforward to explain because it concerns more widely known and unambiguously interpreted situations: recognition of *posture* within Marr's [1982] essential framework for computer vision which underlies a fair share of current research. The present section describes an attempt to augment and extend some aspects of this framework. These aspects relate to the recognition of a particular object from a collection of stored models on the basis of one or multiple (on different resolution levels) configurations of generalized cones and of the characteristics of these configurations, as formalized into indexes by Marr. The extension outlined in this section concerns in particular the recognition of the posture of a human body.

In Marr [1982, pp. 318-325], once the generalized cones and their axes have been identified in an image, recognition (i.e., association of the thus derived description with one of the descriptions in a collection of stored 3-D models) is based on three indexes which primarily express the number, position and orientation of cone axes, while the size of the cone cross sections and other clues play a secondary role. When the descriptions concern rigid objects or non-rigid objects in canonical postures the relative position and orientation of axes is normally sufficient for recognition. In unconventional or otherwise confusing views [Marr 1982, pp. 314–317] it appears that this information is insufficient. In fact, in such cases the whole approach seems to suffer as it is proposed that recognition relies primarily on clues that are only implicit in the derived description. These difficulties contradict the ease which characterizes human perception of unconventional and partial descriptions (provided that there are no insurmountable ambiguities in the grouping of image elements). For example, a still frame of a horse's leg is practically always recognized as such, regardless of its posture and even when contextual clues and textural information are nonexistent. This presumably happens because the form and postures of horse's legs are familiar images, characteristic features of the appearance of a horse and not easily confused with other objects.

Therefore, it is hardly surprising that the collection of 3-D models used by Marr to demonstrate his approach to recognition represent non-rigid forms (bodies of various animal forms) in a single canonical posture. However, each

of these forms is capable of appearing in a wide variety of different postures. By *posture* here we refer to a “position the body assumes in preparation for the next movement. ... Posture therefore involves the concept of balance, muscular co-ordination and adaptation” [Roaf 1977, p.1]. And as posture forms the basis of movement, the starting and the end point of all movements [Åstrand & Rodahl 1986, p.116], by extension we can also include the configurations of body parts that may occur between such balanced postures at certain sampling rates (not necessarily temporal). Therefore, each non-rigid object is preferably described in and by a large number of postures and not a single, canonical one. This number may be reduced to a set of typical postures, each characterized by a different degree of tolerance with respect to the relative position of its components. Still, even this set is too large to be accommodated into a straight collection of models such as that one may infer from Marr [1982, chapter 5]. A good example is the variability of sitting postures for an adult man depending upon the type and form of the seat (chair, armchair, stool, floor, fence, bicycle) and the activities he may be engaged in (eating, writing, reading, watching television, driving a car). Even if we ignore all these differences and consider only symmetrical sitting postures without support of the back, we still end up with not one but three different postures: a forward, a middle and a rear one [Carlsöö 1972, pp.84–89].

Most computer vision systems choose to ignore this variability and choose on the one hand to concentrate on developing sophisticated techniques for appropriately structuring derived descriptions [Minsky 1975] or on elaborating techniques for matching descriptions to models [Winston 1975] and on the other to represent objects by their canonical postures and views, despite that the most frequent argument for that, keeping the model base and the associated rule bases as compact as possible, is contradicted by the size and complexity of the model base of recent recognition systems, even when they concern rigid objects and a limited number of views only [Ettinger 1988]. Furthermore, as computer vision has so far concentrated on rigid motion, investigations on other kinds of motion (elastic motion, fluid motion, turbulent flow) have been sporadic [Heeger & Pentland 1986] and lack a cohesive framework of approach.

While one cannot ignore the significance of canonical views and postures, that is, of patterns which, with respect to the features they contain and the positions they are viewed from are close to the forms we encode and remember patterns of that particular type, as this significance has been amply demonstrated by e.g. Hochberg [1972], it is only fair to say that canonical models result in less dynamic systems which often ignore some of the less obvious but equally critical issues of recognition, such as the activities that may

be associated with which canonical posture. These problems are often accentuated by that in most cases different canonical views and postures are merely cross referenced and classified together in the database of stored models and not merged into integrated comprehensive structures which reflect all canonical or even possible instances. The frame system approach [Minsky 1975] represents a move towards the direction of integration; however, straightforward applications of frame theory support both the integrity of individual frames and their obedience to rather simple central top-down structures.

A primary advantage of canonical model based recognition is computational economy and the related reduction of required memory space. This advantage becomes particularly evident if we consider the alternative of an exhaustive enumeration of all possible views and postures. Marr's answer to that particular problem was to abandon the viewer-centered coordinate system in favour of an *object-centered coordinate system* which makes descriptions independent of the vantage point. This approach resolves the problem in the case of complex rigid objects and also for canonical postures of non-rigid objects but does not account for the variability in posture of non-rigid objects.

The technique of cognitive filtering proposed in the present section is an attempt to accommodate this variability in Marr's essential framework. The stage of recognition in computer vision is normally the point where one is forced to integrate, connect or otherwise relate the data driven processes through which the image is analysed and re-synthesized into a description with the model based descriptions that are stored in memory. This exercise (which we shall henceforth refer to as *integration*, since integration is the most advanced form of relating the two categories of processes and descriptions and their underlying approaches) is a demanding one, even if effectively reduced to simpler problems such as matching simple and isolated measurements (e.g. frame terminals).

In Marr [1982, pp. 321-325] this integration, or rather in Marr's own terms the "interaction between derivation [of a description, the bottom-up component] and recognition" (i.e., matching to a stored model) concentrates on the elaboration and constraining of the derived description. The stored model guides the completion and connection of the derived description but the description is not always contained in the model: the relations between components in the description are usually more detailed and specific than in the stored model. In other words, recognition in Marr is not so much an issue of detecting an instance of a model as a case of *elaboration of a description by reference to a canonical description*.

This position is obviously related to the structure of the

database of stored models, which in Marr is presumed to be a conventional collection of canonical postures: by elaborating a description one is relatively free from the precise posture of the stored model. If, however, we are interested in identifying the kind of activity that characterizes the objects in the description, whether for instance a man is sitting, standing, walking, running or crawling, an expansion of the model base so as to include all possible postures of each model seems inevitable and, in that case, the posture of each model would obviously be more binding and hence turn Marr's recognition technique from elaboration to more straightforward matching with tolerances.

The approach outlined in this section represents a more radical departure from conventional model bases. Stored models are decomposed and collectively reassembled into flexible and non-deterministic structures that are better integrated in the bottom-up derivation of the description of a scene. The flexibility of these structures corresponds with the variability in spatial resolution levels proposed by Witkin [1986], as the proposed technique does not analyse the description derived from an image on some fixed levels but rather utilizes those from a collection of available levels that are best suited to the particular structure of the description.

#### 4.3.2 A multilevel system of cognitive filters

It is common ground in computer vision research concerned with motion that a single still frame is seldom sufficient for the full recognition of a scene. Especially in the case of moving non-rigid objects, such as biped animal forms, more than one temporally distinct frames make recognition of a scene easier.

In such sequences of frames a non-rigid form is identified at different places and in different postures. Therefore, a prerequisite to recognition in this case is preexisting knowledge of constraints that regulate the posture of the form so as to correlate its appearance in different frames and justify the differences. Such knowledge can also be instrumental in the reverse case, that is, in the recognition of a non-rigid form in a single still frame. It can be argued that in a computer vision system, in accordance with the principles of Marr's object-oriented coordinate approach, the stored postures of a model can be substituted by constraints on the relative position of the components of a model. For recognition tasks this information is complementary to that provided by the size and the canonical spatial articulation of the components.

There are two alternative approaches to the encoding of such constraints in conjunction with the other information about a particular form. The first is to attach the movement constraints to one or more canonical postures, in the form of

*tolerances around canonical positions* of the one component relative to another. Recognition remains a top-down process of matching an overall model to the derived description. A second approach is to *decompose* the possible spatial articulation types (i.e., postures) of the components of the model and their movements into postural constraints and then reassemble them all into one multilevel constraint system. In this case recognition is bottom-up and can be thought of as either constructing a model by connecting compatible parts or as identifying a model on the basis of the relationships of its components rather than the mere existence of these components.

The present section presents an example of the latter approach. The choice was mainly due to two reasons. The first is that, since a canonical posture represents both an object and a class of its most characteristic activities, attaching movement constraints to canonical models inevitably emphasizes this class of activities at the cost of all other classes. To overcome this it would be necessary to employ multiple, overlapping models for each object, each model corresponding to a particular class of activities. This, however, runs contrary to the object-centered approach and to the underlying uniqueness criterion of Marr. The second reason relates to the integration of the data driven process of deriving a symbolic description from an image and the model based process of recognition. The section describes an attempt to facilitate this integration by bringing the structure of stored models closer to that of the description and of the process through which the description is derived.

Decomposition of 3-D models is based on the relations between visual perception and the sense of *proprioception*, the sense that supplies information about the position of a part of our bodies in relation to another one and to gravity. Proprioception is not related exclusively with the sense of the static limb position. There are indications that at least one particular class of proprioceptors that is related to the proposed technique, joint proprioceptors, “do not play a crucial role in the sense of the static limb position, but most likely their reports are important in locomotion” [Åstrand & Rodahl 1986, p.80]. The relations between proprioception and movement control and planning are obvious and profound, as we can see from a multiplicity of anatomical and kinesiologic studies [Åstrand & Rodahl 1986; Carlsöö 1972; Espenschade & Eckert 1980; Pick 1984; O’Connell & Gardner 1972]. We might therefore assume that proprioceptive encoding somehow mediates between visual perception and movement. Even in the chaotic reflex-induced movements of infants one can observe direct relations between visual perception and proprioception. Bower [1977, pp.31–31; 1979, pp.304–305] reports that a 6-day old infant mimics his mother who sticks her tongue out

at him by sticking his tongue out in return.

Although studies of the relations between perception and movement have generally concentrated on the influence of the former on the later and left relatively unexplored the reverse case [Prinz & Sanders 1984] and even if we choose not to get involved in the discussion on efference or sensimotor theories of perception [Coren 1981; Matin 1982; Shebiliske 1984; Sheerer 1984; Turvey 1971], we may safely accept that there are certain aspects of perception that are strongly influenced by movement. For example, the segmentation of the visual field into identifiable objects that influences the appreciation of length and direction (and hence relates to most optical illusions) is based on eye movements [Sheerer 1984]. Concern for the influence of movement and action on perception and cognition also appears in other fields: Gombrich [1972] suggests that encoding and recognition of human faces is related to empathic muscular responses by the perceiver.

The suggestion of the present section is that proprioceptive encoding of visually perceived movement may be a fruitful direction for the recognition of posture in computer vision. It can be argued that we associate complete ranges of relative positions and related repertoires of movement with each of the members of a non-rigid form. This type of encoding allows us to relate perceived postures with our own with considerable ease and directness.

Within Marr's bottom-up, object-centered framework, 3-D models and their canonical postures may be substituted by a more comprehensive encoding of the spectrum of possible relative positions for each joint of a model. These descriptions form the basis of low level cognitive filters in our system. Kinesiologic studies provide us not only with these movement constraints for each joint but also with constraints on the relative position of different joints of the same limb or of different limbs. These are investigated on subsequent levels of cognitive filtering which investigate the relative positions of different joints.

The proposed technique involves a multilevel system of cognitive filters which accommodate the various local, regional or global movement constraints and their relationships. Each filter determines in a bandpass manner the acceptability of each joint for particular activities, either simply with respect to the relative position of the components it joins (on low levels) or also with reference to the like acceptability of related joints (on higher levels). Application of cognitive filtering presupposes analysis of the image and derivation of a basic description in terms of generalized cones and their axes but no recognition of the stored model that corresponds to the derived description. This is precisely the objective of cognitive filtering, with the difference that it is aimed at recognizing any posture of a

non-rigid form and not just canonical or typical ones.

Each of the cognitive filters is applicable to different parts of a models and on different resolution levels. As Witkin [1986] has suggested, a strict sequence of specific resolution levels is neither necessary nor necessarily helpful in computer vision. This is particularly true for the proposed system, as different resolution levels, movement repertoires and corresponding cognitive filters are applicable to different forms (models) and to different postures. In general and within the context of the application described here, we can distinguish between three essential categories of cognitive filters which roughly correspond to the answers to the questions, (1) what can this description stand for (e.g. a running man, a walking man, an ape or a chicken), (2) what does the description stand for (e.g. a man), and (3) what is the depicted object doing (e.g. walking). The overall structure of these categories is clearly hierarchical, as one can see from the fact that each question presupposes a reasonable answer to the previous one.

The first category, *low level cognitive filters*, accommodates local movement constraints which apply to each particular component joint. The first task of low level filters is to determine the acceptability of a particular joint as, for instance, a horse's front knee or a human elbow on the basis of their relative position. Cognitive filtering on this level complements the acceptability criteria that can be established on the basis of the size of the joined components. Low level filters also determine the acceptability of the joint with respect to different activities and the corresponding range of positions for that particular joint.

The second category, *intermediate level cognitive filters*, is concerned with local, regional and overall canonical postures. The acceptability of a canonical posture is determined by the categorization of each joint by low level filtering. The placement of canonical postures on the intermediate levels is based on the rationale that such postures are more useful in the recognition of a model type (for instance, a man versus a gorilla) and less useful in the recognition of the particular movement repertory which applies to the derived description (such as a running man versus a walking man). The latter aspect of recognition involves a finer discrimination and classification of the image and has little to do with canonical postures although one might argue that each movement repertory and each activity has its typical or characteristic postures.

The last category, *high level cognitive filters*, connects and coordinates local constraints of low level filters using intermediate levels as guideline, with the purpose of identifying the precise type of activity the perceived object is engaged into. The structure of this category is therefore the most flexible of the three. This flexibility is further enhanced

by that coordination of local constraints is examined on a multiplicity of resolution levels.

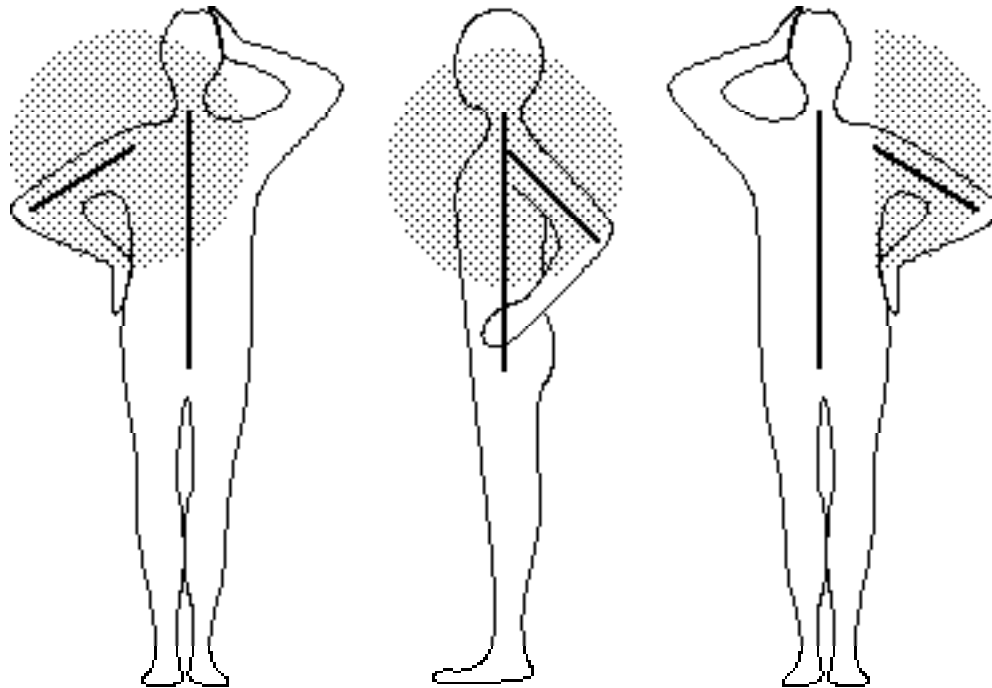
### 4.3.3 Low level cognitive filters

One of the primary types of constraints in the recognition process proposed by Marr are the *adjunct relations* between pairs of axes in a description [Marr 1982, pp.307–309]. These express the position in space of one axis relative to another. Marr suggests that the angles and lengths that occur in an adjunct relation should be represented with variable precision so as to compensate for the variability in the precision of 3-D models. Another reason for including such tolerances is that the adjunct relation between components of a non-rigid form are largely variable.

If we simply allow tolerances around canonical positions we can account for a small number of all possible relative positions, such as variations caused by postural sway. Kinesiologic investigations [even early ones such as Muybridge 1899, 1901] reveal the tremendous variety in posture and in the position of single joints in the body of animal forms. Even if we are concerned only with balanced postures, the fact that stretch reflexes (the reflexes that keep muscles at fixed lengths and thus stabilize the relative position of body parts) allow muscles to be maintained at many different lengths emphasizes the ability to stabilize the body in many different postures [Griffiths 1974, p.459]. This variety cannot be accounted by a single adjunct index, however tolerant it may be. Thankfully, movement at a joint is constrained by a small number of factors common to practically all joints, such as the tension of ligaments, the tension of muscles that are antagonistic to the movement, or the relative position of other parts of the body (for instance, flexion of the hip joint is ultimately limited by the abdomen) [Åstrand & Rodahl 1986, p.284]. This allows us to define the range of possible positions at any joint with high precision.

Within Marr's framework, the range of possible positions of a generalized cone relative to another cone it is joined to in a 3-D model of any resolution level forms a cone (or a planar angle in the case of a uniaxial joint — Figures 4-7 and 4-8) that can be easily defined on the basis of kinematic data. In almost all cases a cone of a 3-D model in Marr [1982] corresponds to an anatomical limb or part of a limb but a joint between cones does not necessarily correspond to an anatomical joint. The best example of this distinction is the shoulder area of the human body: the cooperation between glenohumeral (shoulder)

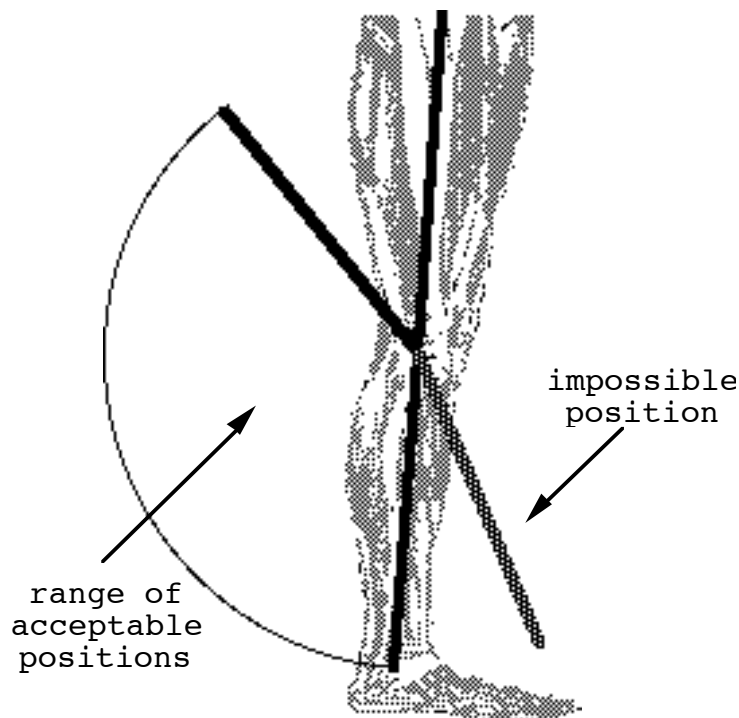
movement and movements of the shoulder girdle is so close that such movements are perceived by the naive observer as a single movement around a single joint [Hay & Reid 1988, p.77]. Marr's 3-D model of the human body reflects this perception rather than its anatomical analysis —and rightly so.



**Figure 4-7** An approximation of the range of possible positions for the shoulder joint in a model of the human body according to Marr [1982]

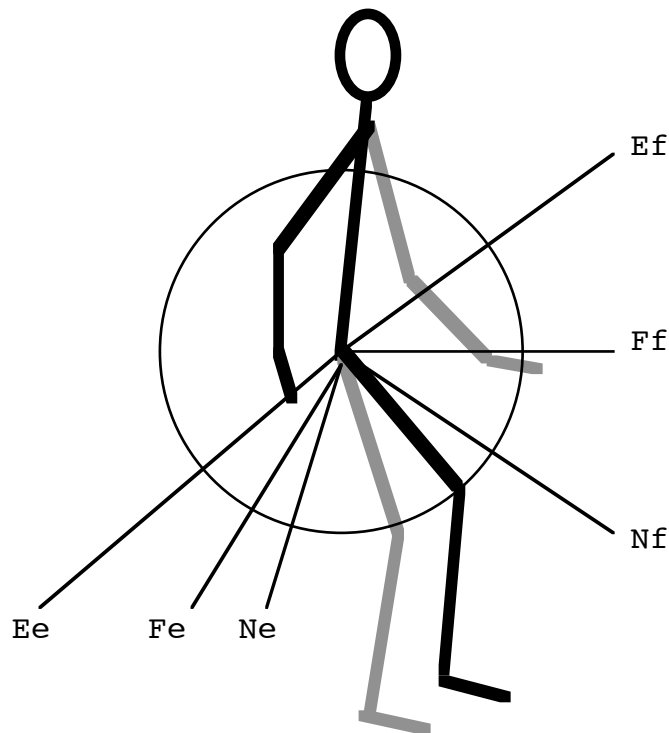
*from left to right: in front of and on the frontal plane (projection on the frontal plane); to the right of and on the sagittal plane (projection on the sagittal plane); behind and on the frontal plane (projection on the rear of the frontal plane)*

Low level cognitive filters accommodate these constraints, that is, they associate different magnitudes and orientations of movement to each generalized cone at each joint. They are local, as each filter concerns only one joint of a model. Their primary purpose is to determine the range of relative positions that are acceptable for each joint. This range, together with the dimensions of the joined components, is helpful in the identification of a form in terms of its parts, even in the case of partial views, as this information is often characteristic of particular models. In an object-centered coordinate system, the range of relative positions, being less susceptible to perspective and other deformations, may be instrumental in recognition. A good example is the thumb of a human hand in comparison to an ape's thumb.



**Figure 4-8** *The approximate range of possible positions at the knee joint*

Low level cognitive filtering does not go so far as to result into an identification of a complete model on the basis of its parts; it is restricted to the investigation of the acceptability of each joint with respect to each model separately. In addition, low level filters allow the identification of possible actions and activities which may cause the relative positions of components at each joint and thus define the postural categories that should be investigated at subsequent levels of cognitive filtering. This identification consists of two parts: firstly, low level filters suggest (often rather strongly) whether a described posture arises from a normal, fringe or exceptional activity (Figure 4-9) and, secondly, they suggest possible variations of a particular model (such as an old person or a soccer player in action) that may correspond to each joint.



**Figure 4-9** The approximate ranges of hip joint position categories on the sagittal plane  
*normal positions may occur between  $N_f$  and  $N_e$ , fringe positions between  $F_f$  and  $F_e$  and exceptional positions between  $E_f$  and  $E_e$*

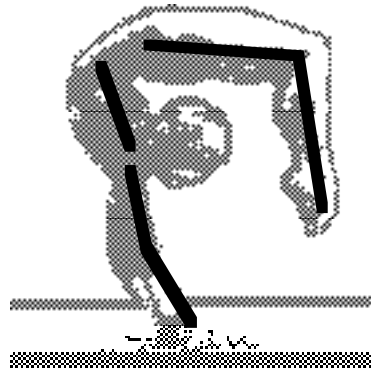
*Normal* encompasses commonplace everyday activities, such as walking or sitting. The relative position at a joint for normal activities is usually characterized by a rather limited range with respect to what is possible in general. For example, in walking the relative positions at the hip joint cover only a subset of the range allowed by the involved ligaments and muscles. The reason is that, in walking, flexion of the knee joint is limited by the length of the muscles on the back of the thigh because the knee is at or near the extended position. *Fringe* refers to activities which are characterized by a wider spectrum of relative positions that come close to or even reach the limits of movement at a joint. Sports are examples of fringe activities. Running causes a wider range of positions at the hip joint than walking; in fact, flexion and extension in running reach the limits of movement by muscle contraction only, that is, without the application of an external force. This wider range is allowed by that in running the knee is flexed far more than in walking.

*Exceptional* refers to these activities which usually elicit the observer's surprise and amazement because they violate the constraints for the relative positions of components at one or more joints. We should distinguish between two different kinds of constraint violations. The first concerns violations of strong constraints and denotes impossible postures. For example, in the case of the human

knee, the patella (kneecap) poses insurmountable objections for certain positions of the lower leg relative to the thigh (Figure 4-8). If such positions are found in an image of a presumed human body, then some error must have occurred: either the description is not of a human body or the parts of the image that are supposed to form the leg do not belong to the human body or even some very unfortunate accident has happened to a human being.

On the other hand, there are cases where a position may fall outside the normal range and still be possible. Here again we should distinguish between two subcategories of exceptional positions. In the one a joint is encountered in positions that go beyond the limits of fringe positions but still remain within what is normal for anatomy as a result of an exceptional muscular contraction or more usually due to the application of an external force. For example, flexion of the hip joint can be increased further than in fringe positions by pulling the knee towards the body. In this case flexion is actually stopped by the thigh pressing against the abdomen (position E<sub>f</sub> in Figure 4-9). In other words, external forces may cause movement beyond that can be achieved by muscular contraction.

The second subcategory of acceptable even though highly improbable exceptional positions concerns movement beyond what is considered normal in anatomy and kinesiology. A contortionist's body provides us with numerous examples of such improbable positions made feasible by acquired skills, that is, through exercise that has increased the length of muscles and hence the flexibility of the joint [Åstrand & Rodahl 1986, p.284] (Figure 4-10). The identification of such exceptional activities usually requires augmentation of the movement repertoires of one or more joints and controlled relaxation of corresponding constraints. The results of these modifications on the low level is a minor enlargement of the range of positions not only at the joints where some violation of normal constraints has been observed but also at other, causally related joints. These modifications are particularly significant for higher levels of cognitive filtering because, while locally they are simply an issue of magnitude, the overall postures related to exceptional activities are usually far removed from those that correspond to the movement repertoires of, for instance, a walking man or a basketball player.



**Figure 4-10** *The exceptional joint positions in the body of a contortionist*

*acquired extreme flexibility of the lumbar vertebrae allows hyperextension well beyond the acceptable limits for the hip area joints in a model of the human body according to Marr (after Hay & Reid [1988])*

On the low levels of cognitive filtering such modifications can be further exploited by employing different ranges of positions for different instances of a model (or, if you prefer, for each member of a class of models). For example, we may distinguish between the movement repertoires of normal adults, the augmented repertoires of ballet dancers and the restricted repertoires of old people. Each of these sub-models is characterized by different ranges of relative positions for each joint. For practically any joint the range of old people is a subset to that of normal adults, which in turn is a subset to that of ballet dancers. In such cases, the distinction between the aforementioned two kinds of violations of movement constraints at a joint can be particularly instrumental in recognition. Suppose that at a certain point in the process of recognition expectations concentrate around the model of old people. Then some signs of unusual agility, that is, relative positions at a joint that go beyond the range of this sub-model but are still within the corresponding range of other sub-models of the same type can trigger a reappraisal of all preceding components of recognition and a more accurate calibration of expectations which might lead the recognition of, say, a normal adult or a ballet dancer disguised as an old man. Such violations are entirely different from these which reveal that recognition had so far followed a wrong path, such as when from a particular movement of what we thought to be the knee we discover that the figure we see behind the bushes is not a man but a deer.

Each joint position category includes a small number of crucial exceptions. For example, most sitting postures are related to normal everyday activities but extrinsic factors (such as the form or size of the seat) cause joint positions that may range from fringe to exceptional. Low level cognitive filtering does not attempt to cover such exceptions; this is a task of intermediate and high level filters where the exceptions are identified through the correlation of joint

positions. An alternative which would allow us to identify the exceptions in low level cognitive filtering would be to distinguish between different kinds of movement which are organized in different relation to the environment (such as postural, locomotor transfer and manipulation of external objects). This would require further information about the scene depicted in the image than that is input to the cognitive filtering system presented here.

#### 4.3.4 Intermediate level cognitive filters

Once the acceptability of a description has been examined on the level of local conditions at each joint, we can proceed to the overall recognition of a form. This level of recognition is addressed by intermediate level cognitive filters. These attempt to recognize the derived description by comparing it to descriptions of canonical postures of specific objects.

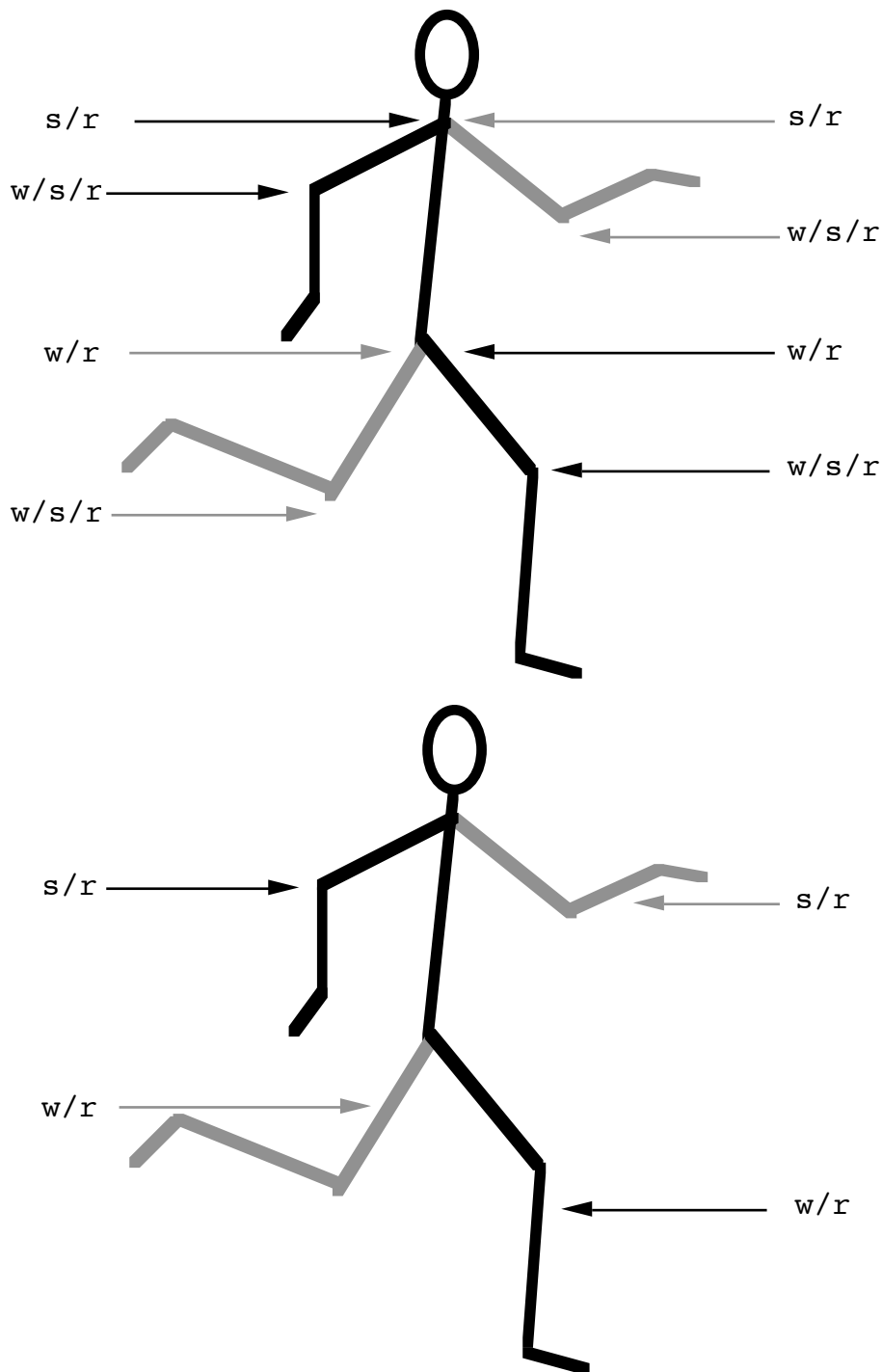
The purpose and structure of this level is similar to the recognition process on the basis of specificity, adjunct and parent indexes, as described by Marr [1982, pp.318–321]. The main difference lies in that for each model there can be more than one canonical postures, defined in terms of alternative configurations of joint positions for the configuration of generalized cones that describes the model in general. For example, there can be a model of a standing, a sitting, a running and a crawling. Each of these postures relates to different movement repertoires which correspond to different though overlapping ranges of relative positions at each joint. Therefore, recognition does not depend solely on the specificity and parent indexes of Marr but also on the results of low level cognitive filtering. Low level filters provide us with probabilities as to which model a joint may belong to. Constraint relaxation as in Waltz's procedure [Waltz 1975] can be used to relate connected joints and thereby assist or verify recognition (Figure 4-11).

On the intermediate level the range of relative positions for the components of each joint can be considered as a kind of tolerances around canonical positions but only in the sense that it allows variability in the definition of a canonical position or rather the abstraction of all possible types of postures into a relatively small number of general categories which can be considered as canonical postures. The description of an overall canonical posture in terms of its constituent parts, that is, the canonical positions of components at each joint, with variable precision and within an object-centered coordinate system has one primary advantage: it allows overall recognition (i.e., correlation of local constraints) to be performed hierarchically, on a variety of levels. The first level is that of directly connected joints

which together describe a distinct part of the model, a *region*, which usually corresponds to a limb, such as a leg or an arm. Then follow different levels of connected an/or related regions, such as the arms, the legs, the right arm and leg or the right arm and left leg. These combinations are characteristic of specific postures and hence facilitate recognition even in the case of partial descriptions. The number of levels and the definition of regions depends on the detail of the image, the specificity of the derived description and of the models, and other related factors, as in Witkin [1986].

In the case of partial or unconventional descriptions the use of cognitive filtering can assist in bypassing problems such as those mentioned in the case of a frontal view of a horse as opposed to a side view [Marr 1982, pp.314–316]. These include not only that the derived description covers only part of the model but also that the description of the torso component is entirely different from that in the stored model as a result of foreshortening. Marr suggests two approaches for dealing with such problems. The first is the use of partial descriptions as models, an approach which admittedly weakens the representation, and the other is to exploit the constraints of the components which can be recognized in an unconventional view.

Intermediate level cognitive filtering can be considered as a fusion of the two approaches. It relies on the systematic propagation of constraints from each joint and, being bottom-up, proceeds from part to whole and not “from the general to the specific”, as advocated by Marr [1982, p.321] and hence provides a succession of partial models (regions and region configurations) closely correlated with the whole. In the case of a frontal view of a horse, three distinct regions of the whole model are recognizable: the two forelegs and the head with the neck. The forelegs have quite concrete relations between them in any posture and form a coherent region which in turn has very specific relations with the region consisting of the head and neck. This information is normally sufficient to suggest that the image may be of a horse and that the other visible component in the image may be the torso of a horse and perhaps even that divergence from the canonical form of a horse’s torso may be due to some awkward view.



**Figure 4-11** *An approximation of intermediate level cognitive filtering in posture recognition*

*top: joint positions are input with indications of their acceptability as parts of a walking (w), a running (r) and a sitting (s) posture; bottom: correlation of joint position categories on a regional level reduces the acceptability of each region with respect to these categories; further correlation obviously leads to the identification of a single category (running) for the whole body in the particular case*

One final point about intermediate level cognitive filters concerns the use of multiple models to describe a single object. As we have seen, this multiplicity contravenes the uniqueness criterion which underlies Marr's object-centered approach to the representation of stored models. In

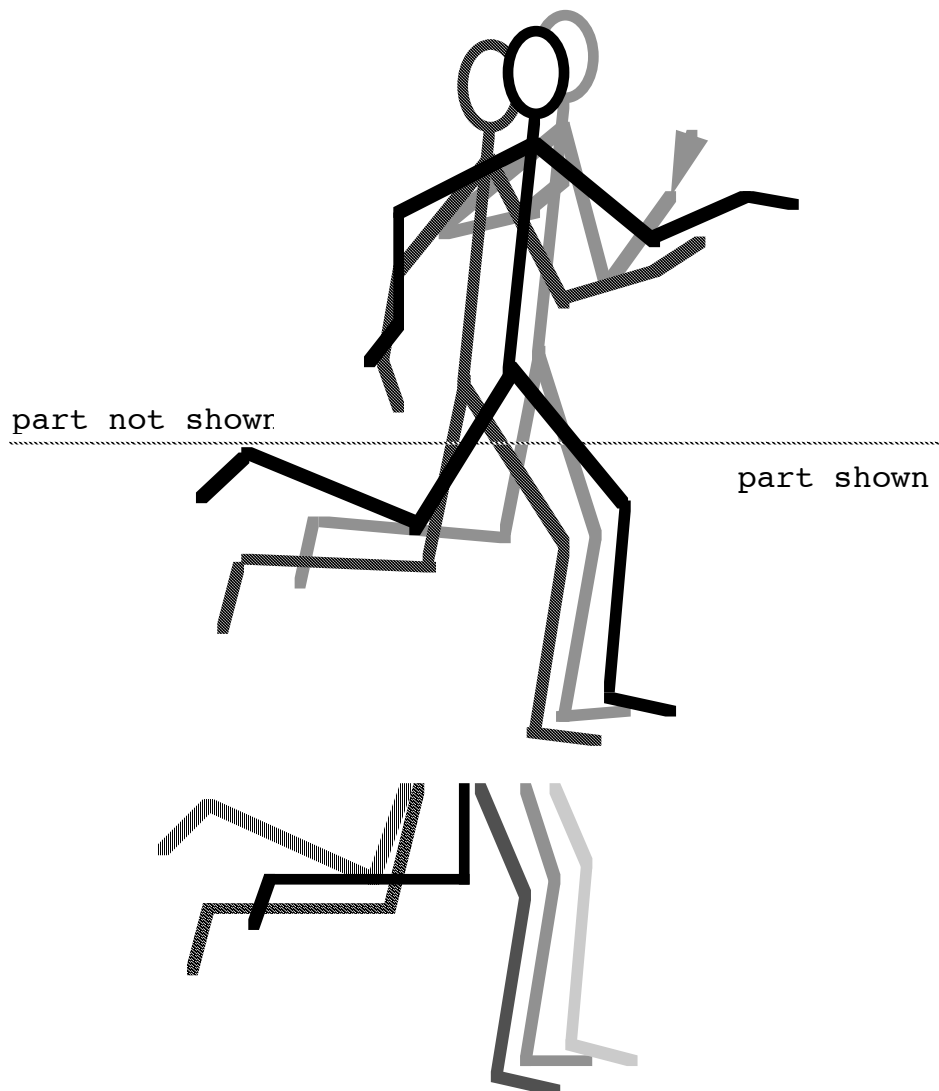
our case, the acceptance of more than one models for the same object relates to the nature of canonical and typical postures in perception. Such postures can be considered as a kind of *mnemonics* for different categories of actions and activities. An example is provided by photographs of sports events which succeed in “freezing” activity and motion into characteristic postures which implicitly describe intricate patterns of movement. The number and extent of disparities between canonical postures necessitates the use of a separate model for each posture. These models naturally cannot account for every possible posture of an object but can cover some of the more frequent ones and restrict the number of possible matches in cases where no single model can be recognized in a derived description. In fact one could say that this has been one of the cornerstones in most model based approaches: even a limited number of variations of a model can be an adequate data base for some essential recognition tasks.

#### 4.3.5 High level cognitive filters

High level cognitive filters represent the additional elaboration in recognition that is allowed by the proposed technique. This elaboration is on the one hand directed towards identifying the particular class of activities that are related to the posture of the recognized non-rigid forms and on the other offers an additional verification of the results of the preceding stages of description and recognition. High level cognitive filtering dispenses with canonical views and postures, which, although instrumental as a kind of index in a classification structure and also for the recall and recognition of general categories, do not ultimately suffice by themselves as the memory component of a perceptual process.

High level cognitive filtering can be described as the correlation and coordination of local constraints at each joint in a hierarchical manner, starting with connected joints that form concrete regions and then proceeding through connected or otherwise related regions to the recognition of the overall posture. As with intermediate level cognitive filtering, these hierarchies allow for the recognition of partial descriptions as well as for the correlation of partial descriptions. For example, consider an image of a number of human legs where the connections between each pair of legs are not shown, e.g. of a pack of middle-distance runners or a group of soccer players, an images that is not an unpopular subject in photography (Figure 4-12). Any normal person confronted with such an image has no difficulty in identifying each and every leg with accuracy nor in identifying pairs of legs as belonging to one person (with the

occasional mistake that is almost always immediately corrected).



**Figure 4-12** An image depicting an unconnected number of legs  
top: the whole scene; bottom: the connected regions that are identifiable  
in the part shown

Intermediate level filters can establish that all depicted objects are human legs. However, except for clues of orientation, occlusion and size — which can be very scarce and inconclusive — we cannot determine which pair of legs belongs to one and the same person. High level filters attempt to fill that gap by correlating the posture of each leg with that of the others and with the overall posture. From the previous levels of cognitive filtering we already know whether each leg corresponds to a running, a standing or a stumbling figure and also the specific position at each joint. By combining the three basic position categories (normal, fringe and exceptional) used in low and intermediate level cognitive filtering with anatomic / kinesiological classifications which relate to the mechanics of the muscular and skeletal systems, so that every resulting subdivision of

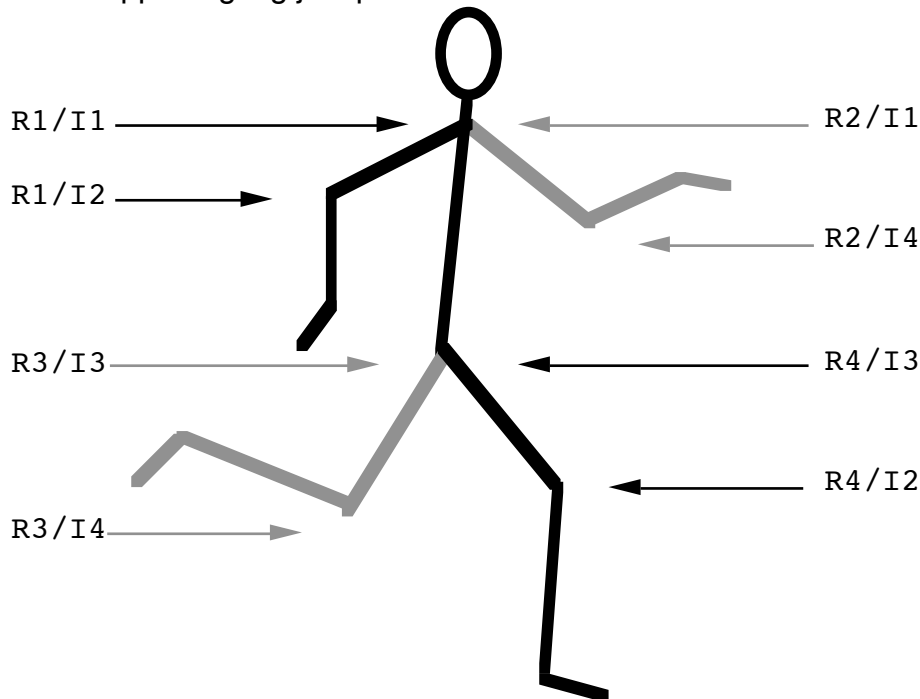
the range of possible positions at each joint corresponds to a single category of each constituent classification (such as exceptional flexion, normal extension or fringe adduction) we arrive at a finer categorization of joint positions that connects each type of activities with the underlying kinetic systems that determine the range of each category.

This finer categorization of joint positions is instrumental for the correlation of joints whether they belong to the same limb or not. The correlation of adjacent joints is justified by the structural connections (muscular and skeletal) of joints of the same limb: “when any of the three joints of a limb is flexed, the other two joints also flex. When one is extended, the others also extend” [O’Connell & Gardner 1972, p.209]. The correlation of interlimb joints is justified by more functional connections. When a standing person performs a precision movement with the fingers, a subconscious ‘background activity’ of the leg, trunk and arm muscles facilitates the conscious movement by stabilizing the hand [Åstrand & Rodahl 1986, pp.95–96]. This background activity consists of motor responses evoked by mutual facilitation of interlimb joint proprioceptors [O’Connell & Gardner 1972, p.210].

Such movement constraints have been rigorously exploited in kinetics so as to explain and improve coordination of body movements in performing certain tasks, such as in sports. These analyses are often based on representations of the body as a system of connected axes. These include the axes of generalized cones in the representation of Marr and also axes which connect endpoints of cone axes, such as the hip and shoulder axes. The proposed technique of cognitive filtering does not make explicit use of the latter category of axes, exactly because they are not encountered in Marr’s system. Movement constraints represented by these axes are covered by the relations between different regions or configurations of regions.

On the basis of the above and the abundance of existing kinematic data it is possible to correlate the position of different joints on a regional level and then correlate regions up to the level of the description or the whole model. The end result of high level cognitive filtering is (a) an unambiguous recognition of the model and (b) a tentative recognition of the precise actions and activities that correspond to the description. This allows us to identify whether, for instance, the posture of a soccer player corresponds to simply running or kicking the ball and, in the latter case, whether he is about to kick the ball, is precisely at this moment kicking it or has already kicked it. This identification is very useful in connecting partial descriptions, as those

in Figure 4-12, since the one leg of the same person cannot be kicking a ball while the other has the posture of the supporting leg just prior to the kick.



**Figure 4-13** An approximation of the joint correlation system employed in high level cognitive filtering  
*joint positions are initially correlated within four distinct regions,  $R_1$ ,  $R_2$ ,  $R_3$  and  $R_4$ ; these regions are correlated pairwise on the basis of positions at the joints where they link with the torso (e.g.  $R_1$  and  $R_2$  are correlated on the basis of the joints marked  $I_1$ , and  $R_3$  and  $R_4$  on the basis of  $I_3$ ) and on the basis of the positions of equivalent joints (e.g.  $R_1$  and  $R_4$  on the basis of  $I_2$ , and  $R_2$  and  $R_3$  on the basis of  $I_4$ )*

The precise structure and sequence of high level cognitive filters depends very much on the concerns of the recognition framework they are implemented within, that is, the detail of the image, the specificity of the description derived from it, and the specificity and variability of the models. In general, high level cognitive filtering operates in the following manner: First, regions suggested by the intermediate levels are verified and the corresponding postures and activities are investigated in more detail, this time with reference not to canonical postures but to *repertories of movements* represented by configurations of joint position ranges. This stage results into a limited number of possible movement types for each region (Figure 4-13).

A recursive correlation of these regions is attempted on the basis of two classes of criteria (in the case of the human body): (a) the position of the joint that connects the region to the central (torso) component or region, such as the position at the hip or the shoulder, and (b) the positions of functionally related interlimb pairs of joints, such as the knee and the elbow. For instance, in running forward when the left

hip and knee reach maximal extension, the left elbow is at maximal flexion, the right hip and knee at maximal flexion and the right elbow at minimal flexion. Joint positions oscillate in a consistent and regular manner between these and the opposite limits for this particular type of activity, thus defining very concrete and specific general patterns of interlimb joint position relationships. The correlation of regions according to these patterns results into a positive identification of a single repertory of movements, that is, of a particular type of activities, for a particular model.

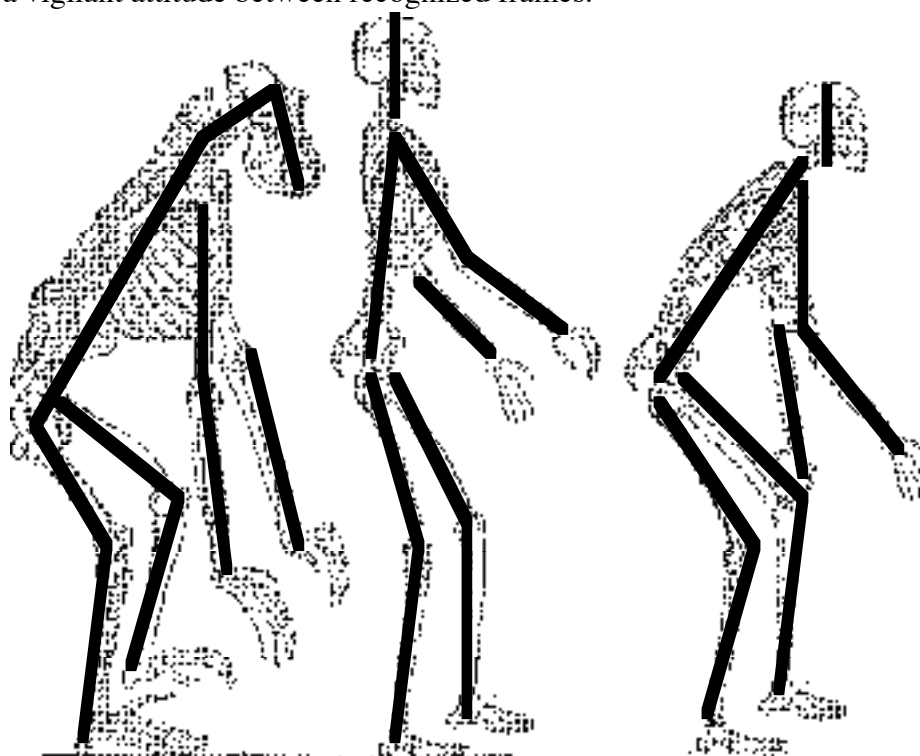
#### 4.3.6 Conclusions

The proposed technique of cognitive filtering suggests that recognition of a non-rigid form can be facilitated by the association of movement constraints with each joint. These movement constraints represent the variety of postures the form may assume in the course of different activities and thus complement the recognition process proposed by Marr by allowing also recognition of the posture of the form depicted in an image.

This elaborate recognition of human posture is obviously of little significance in cases where a derived description matches perfectly or roughly a stored model of a particular posture. However, these cases represent a minority among the possible and probable appearances of a human figure or any other non-rigid form and in particular of those capable of locomotion. The variations of each canonical or typical posture are so many that it is not possible to recognize them all with reference to a limited number of stored models. Storing a vast number of models does not resolve the problem automatically because management of a huge database can be a major drawback for any system and also because however large the number of stored models there would still be variations, however few, around each of them. By verifying the possibility of a posture through the use of specific local constraints at each joint and general coordination constraints which connect the joints we can dispense with all but the most essential canonical models and increase variability and flexibility in the recognition of non-rigid forms. This approach to recognition offers the additional advantages of allowing the formation of very specific expectations for the next (and previous) actions of a recognized form, of unambiguous partial recognition on the basis of complete models, and of rather effortless differentiations in the movement

repertoires due to occasional augmentation with unrelated action types, such as when the winner in a race raises his hands while crossing the finish line or when a marathon runner has a drink while running).

The emphasis on precision and detail is considered the main problem for applying most motion analysis approaches on real imagery and in real time. The solutions proposed so far have concentrated on abstract representations of object shape, scene layout and object / observer motion [Thomson & Kearney 1986]. The ability to recognize the posture of a non-rigid form in a single still frame, suggests an alternative approach: the analysis of a scene at temporal intervals which allow complete and precise recognition of a still frame. If recognition of this static image includes such dynamic aspects as posture and associated movement repertoires it is possible to lay out detailed and highly plausible expectations about the movement and related actions of the recognized forms. These expectations can fill the interval between the analysis of a frame and that of the next one. For example, if the system supplies information to an automated guarding system, posture is often sufficient for predicting an approaching person's intentions and trajectory and thus elicit an appropriate response from the guarding system and ensure a vigilant attitude between recognized frames.



**Figure 4-14 Mimicry**  
*from left to right: an ape; a man; a man mimicking an ape*

Another aspect of cognitive filtering is that it is essentially a bottom-up recognition technique whereby a model is in a sense synthesized from compatible parts encountered in a description instead of being juxtaposed

(matched) to a derived description. This does not necessarily improve the efficiency of recognition but certainly allows a smoother transition from the bottom-up descriptive process to recognition. One of the advantages of this is that it may increase the comprehensiveness of each stage of the descriptive process by the integration of recognition elements. The most significant advantage of this approach is that it allows both greater accuracy and greater variability in recognition through the ability to combine aspects of different models. This is perhaps best exemplified by the ability to recognize cases of mimicry: when a man mimics the movements of an ape the dimensions of his limbs and hence of the generalized cones of his description do not change; it is the joint positions that copy the movement repertoires of apes thereby achieving the perceived 'apeness' (Figure 4-14). Another, more sophisticated example of this is Big Jim's hallucination scene in Chaplin's *The Golddigger*: when Chaplin appears as a chicken the dimensions of his body are completely deformed by the chicken suit. When he resumes his normal appearance this deformation disappears (although Chaplin's arms remain close to his sides) but his movements are still so reminiscent of a giant chicken that the transition of Big Jim's mind from hallucination to normal perception is very plausible.

## 4.4 Cognitive filtering in chap

### 4.4.1 The purpose of cognitive filtering in chap

The previous section introduced the basic structure of cognitive filtering through the example of posture recognition within Marr's computer vision approach. This example illustrated that even in cases where recognition is strongly related to matching to complete models, decomposition of the models and reorganization of their components into a bottom-up structure may be advantageous for elaborate recognition tasks. Also recognition of posture has strong similarities with the investigation of well-formedness in chap. Both are based on a proscriptive system of local and global constraints. In both acceptability of the whole depends upon the correlation of acceptable parts (regions and location groups). This correlation facilitates understanding of general factors that are only implicit in the description or the model but nevertheless quintessential elements of the recognition framework: recognition of posture facilitates identification of the underlying system of

kinetic forces, while the investigation of the acceptability of an architectural plan as a classical composition is also an investigation of the ways different schemata of the classical canon may correlate. Furthermore, both recognition systems are superimposed on descriptions that are essentially sufficient for the representation of the image and the basic recognition of a model: in both cases cognitive filtering adds more aspects to the recognition of these descriptions in that it provides more detail to the models recognized in the description, as well as in that it allows better correlation between models.

In the context of posture recognition cognitive filtering was employed to address the tremendous variety of appearances a single form (model) may adopt and thus determine the acceptability of a description in a more comprehensive and flexible manner. Both comprehensiveness and flexibility are supported by the *combinatorial* structure of cognitive filtering which in posture recognition operates by evaluating the parts of the description independently and in relation to each other and the whole. In the case of architectural plans this combinatorial structure allows first of all a distinction between the well-formedness of the parts of an architectural plan and the well-formedness of their correlation, that is, the well-formedness of the whole.

To give an example more to the taste of architects (even though it contaminates the purely formal framework of CHQP with functional concerns), let us suppose we are given a complete and detailed description of a building and its immediate context, including its orientation, and we are asked to give a preliminary evaluation with respect to natural lighting and solar gain. If the latitude of the place is not known then evaluation is not possible. Or rather, what most of us would do, especially if we failed to notice the omission, would be to assume that the latitude is approximately that of the place we currently happen to be or we are accustomed to. A person living in the northern hemisphere would accordingly evaluate the building as if the south side had maximal exposure to sunlight. This evaluation would be completely erroneous if the building lies somewhere in the southern hemisphere. Nevertheless, the evaluation could be very easily (on the conceptual level at least) adapted to such a change by a simple reversal of a single reference frame and without any other modification to the structure of the evaluation.

The magnitude of the evaluation error and the contrasting ease with which it can be corrected suggest a separation into two components, the purely descriptive part of the evaluation and its interpretation, a separation that is evident in the structure of CHQP. Both components contain elements of recognition, each of different, complementary

aspects and levels. The basic description of a building with respect to natural lighting and solar gain in our example remains the same whether it is in the southern or the northern hemisphere. What changes in the evaluation structure is a number of elements related to the interpretation of this description, in our case the trajectories of the sun, that is, the reference frame that attaches natural lighting and solar gain values to the parts of the description with respect to their orientation.

Underlying in CHQP is the assumption that every form of recognition, from any point of view, follows the same pattern of an analytic description that is specific to the recognition task in general and of a superimposed interpretation that is specific to the particular purposes of recognition. A description of an architectural plan in CHQP is as complete as it needs to be so as not to give rise to ambiguities concerning the general form of the plan and its articulation. Cognitive filtering complements this description by relating the instance to the canon and thus allowing its categorization and formal evaluation. In general, cognitive filtering attempts to account for the ultimate differences in the perception of of the image, that is, to explain why a description which remains structurally the same (and is recognized as the same by different perceivers) invokes different interpretations, even under similar conditions.

A corollary of the distinction between description and interpretation is that the interpretation, that is in the case of CHQP, the investigation of well-formedness, applies differently to the parts of a description and to the overall structure of the description as determined by the correlation of the parts. The well-formedness of parts is more an issue for the derivation of the description and is largely covered by the grouping constraints of Chapter 3. The remaining aspects of part well-formedness concern the relations between parts and between parts and the structure of the whole are an integral part of cognitive filtering. As for the well-formedness of the whole architectural plan, this can take the form of an evaluation of the formal quality of each plan with respect to other plans on an absolute or relative scale. Such evaluation is no concern of CHQP and, since all plans included in it are acknowledged classical compositions, there is no point in attempting to verify their conformity to the classical canon in a computer vision system (a verification that can be performed through quite simple devices such as the gratings of Figure 4-5). The investigation of well-formedness in CHQP goes beyond this point by attempting to *analyse* the architectural plans with respect to the classical canon. This analysis is implicit in the descriptive process of the previous chapters. In the present chapter the analysis becomes explicit through cognitive filtering and takes the form of evaluation of the alternative descriptions returned by

the process described in the previous chapters with respect to the classical canon. Evaluation determines on the one hand the preference order for the alternative descriptions and on the other attempts to make explicit the precise forms of the various classical schemata in the particular plan.

#### 4.4.2 Cognitive filtering on the level of individual locations: proportions

Investigation of well-formedness in an architectural plan starts with taxis, the framework of classical composition. On the level of locations cognitive filtering of taxis in CHAP can be distinguished by whether it concerns rectangular locations or locations of complex shape. In the former case the only issue that needs to be considered is the *proportions* of the location while in the latter the shape of the location must also support tripartition. Locations of complex shape are considered with respect to tripartition in subsection 4.4.3 either as solitary locations or as members of groups which also include rectangular locations.

Proportionalities have been quite extensively investigated in architectural theory and history as a result of the explicitness and the extent of their application in classical architecture. To take only studies on Palladio into consideration (cfr. subsection 4.2.2) Wittkower [1952, p.66] suggests that relationships of proportion are a primary reason for the universality of the 5 x 3 pattern in Palladian villa plans, while Ackerman [1977, pp.160–162] considers tripartition as a formal device for the coordination of proportionalities. Also a system that generates Palladian plans in the manner of rectangular dissections [Freedman 1987] directly derives from the proportional systems employed by Palladio [Howard & Lonair 1982].

Despite the consideration they receive in studies of classical architecture, proportions do not enter CHAP as an independent module or as an explicit aspect of other modules, as in CHAP architectural plans are treated more or less as dimensionless patterns, in the sense of March [1976] and of rectangular arrangements. The reasons for the omission of proportional considerations in CHAP are two: the relative weakness of proportionalities and their relations with the levels of the classical canon in the system of Tzonis & Lefaivre [1986].

Proportional relations are often weaker than other aspects of classical composition, especially on the level of individual locations or groups of locations, possibly to the multiplicity of possible acceptable proportions or to their local character as compared to global schemata such as

tripartition and the grid. Strict applications of proportional systems that dominate other formal or functional aspects of design have often caused severe controversy concerning not only the functionality and the scale of the resulting composition but also its overall formal quality.

Proportions can be approached as a joint product of symmetry and taxis, a correlation of metric patterns (rhythm), architectural figures, grid and tripartition. This correlation is obviously performed on a variety of levels, as defined by the subdivisions of genera and also by the levels of grouping described in Chapter 3. Given that such correlations are actually parts of the grouping and cognitive filtering procedures (where they obviously enter *chap*), there seems to be no reason for including proportionalities into an apart cognitive filtering level.

Explicit use of proportional relationships is considered only as a potential extension of the present version of *chap*, one that may facilitate recognition of the expressions of the formal schemata considered in this version. For example, the proportions of the locations that comprise a group can suggest possible operations of fusion, addition, embedding or deletion that have caused the differences between its appearance and that of the part of the 3 x 3 pattern it occupies and hence facilitate recognition of tripartition in the group. As this can also be achieved without any use of proportions we may suggest that proportions may improve the efficiency of recognition and not its effectiveness.

#### 4.4.3 Cognitive filtering on the level of location groups: tripartition

Beyond the level of individual locations the rectangular grid of the plans of *chap* plans restricts investigation of taxis to that of tripartition. Tripartition is investigated locally and globally, that is, within each location group of the descriptions derived through the process of Chapter 3 and in the whole architectural plan as described by the configuration of these groups.

Tripartition in a location group simply means subdivision into three parts one of which occupies a central position. Therefore, investigation of tripartition in simple (first level — cfr. subsection 3.4.1) location groups is facilitated by the identification of its centre. This centre may be a single location or a pair of locations. Complex groups (i.e., groups produced by the combination of simple groups on the second level of grouping — cfr. subsection 3.4.2) may have either a single centre or more than one centres, generally depending upon the number of directions of

articulation that characterize the group. Also, in some cases the centre is ambiguous, as in groups which comprise only two locations.

In a group characterized by a single direction of articulation and a single relationship or two parallel relationships (i.e., a simple group) we can distinguish between the following cases with respect to tripartition:

- groups with three locations
- groups with more than three locations
- groups comprising only one location (solitary locations) of complex shape

In simple groups with three locations the central location is the centre of the group and each location represents a subdivision of the group. The schema of tripartition is represented by the extension of the edges of the central location until they meet the perimeter of the group (Figure 4-15a).

This representation of the schema of tripartition is derived from spatial reasoning and in particular from CITYTOUR [André et al 1987], a system for the description of image sequences in natural languages. A recent comprehensive and systematic overview of spatial reasoning studies as related to projective prepositions (which includes among others Herskovits [1986], Klein [1982, 1983], Miller & Johnson-Laird [1976], Talmy [1983]), that is, prepositions which “convey information about the direction in which object is located with respect to the other” [Retz-Schmidt 1988, p.95], provides useful insights on the nature of an appropriate representation of tripartition in CHQP as it reveals a relatively broad consent to the use of most projective prepositions and a general acceptance of four essential prepositions in the horizontal plane: in front of, behind, to the left of, and to the right of — a scheme which has striking correspondences with the schema of tripartition.

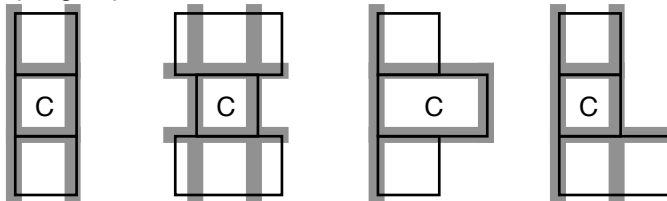
In rough lines we could describe projective spatial reasoning as a system of three essential elements. The first is a *reference structure* which in the horizontal plane consists of four general directions which may segment the world in a number of ways. A Cartesian coordinate system is such a reference structure. The second element is an *origin*, a *reference point* for the reference structure. The third element is the *viewpoint* of the speaker or viewer, normally expressed with relation to the reference point. For CHQP only the first two elements are required, since an architectural plan is a static spatial configuration.

The particular implementation of this scheme in CITYTOUR consists of a subdivision of the horizontal plane into four overlapping half-planes, a front, a back, a left and a right half-plane, which are defined with respect to a

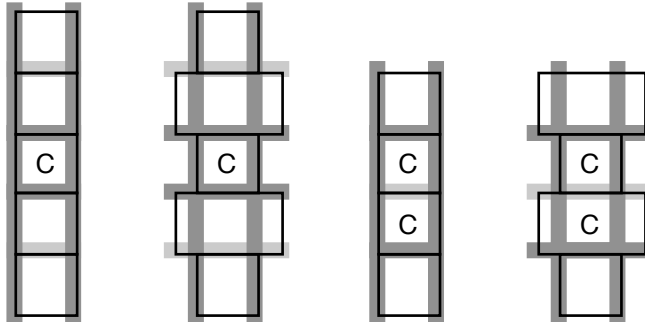
reference object. The half-planes are defined by the edges of a delineative rectangle around the reference object determined by the position of the observer, the reference object itself and other secondary factors (Figure 4-16). A superimposition of the four half-planes yields a structure essentially identical to the 3 x 3 pattern, while allowing a wider scope and hence greater flexibility in the definition of the four projective directions with respect to the tripartition centre.

This scheme is considered preferable to other ways of subdividing an image with respect to a number of directions and one or more reference objects which have been employed in character recognition, spatial reasoning or document understanding, such as loci features [Tsukumo & Asai 1986] or networks connecting landmarks [Davis 1983; also cfr. the hierarchy of space description in Kuipers & Levitt 1988] because it supports better the representation of a superimposed abstract coordination pattern which organizes not only the perception of an image but also the depicted entities.

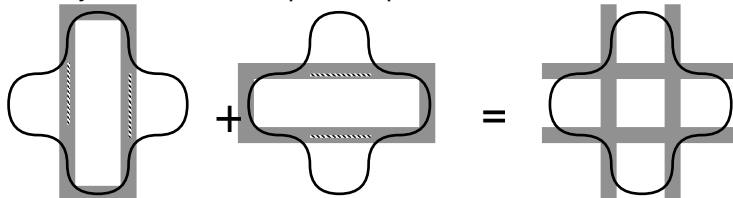
a: simple groups with three locations



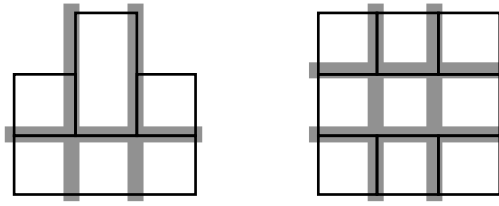
b: simple groups with more than three locations



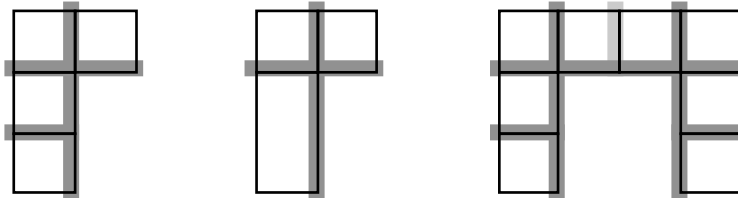
d: solitary locations of complex shape



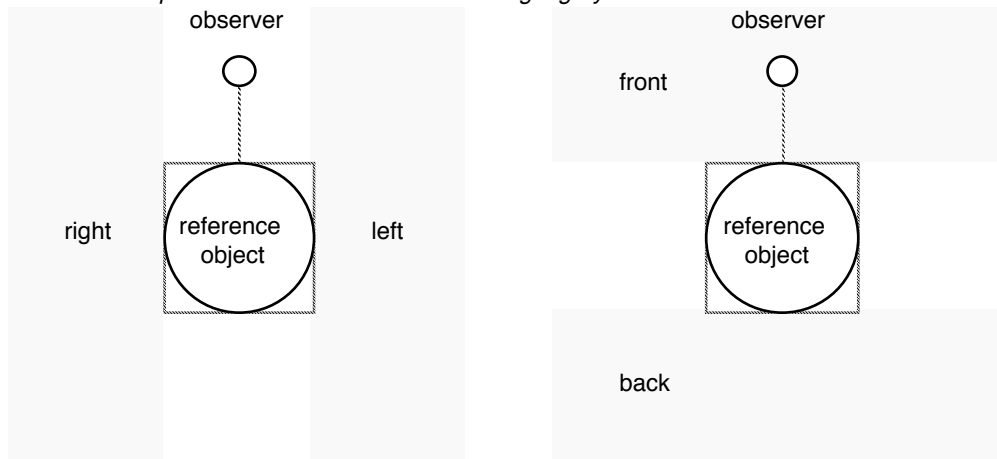
e: complex groups that are subdivided into non-overlapping subgroups



f: complex groups that are not subdivided into non-overlapping subgroups



**Figure 4-15** Representation of tripartition in location groups  
 location boundaries are shown as thin black lines; central locations are marked with a 'C'; top level tripartition lines (i.e., those that subdivide the whole group in three parts) are shown as thick dark gray lines; secondary tripartition lines are shown as thick light gray lines



**Figure 4-16** The CITYTOUR reference frame [André et al 1987]

In simple groups with more than three locations we can distinguish between two subcases: groups with an odd number of locations and groups with an even number of locations. The difference between the two lies in that in the former the centre comprises one location while in the latter of two locations. In these groups the schema of tripartition is represented by the pattern defined by (a) top-level tripartition lines, i.e., the extension of the edges of the central location (or of the rectangle circumscribed by the central pair of locations) until they meet the perimeter of the group, and (b) secondary tripartition lines within each the three subdivisions defined by the tripartition of the whole (Figure 4-15b). The latter are derived by treating each subdivision of the group as if it were an autonomous location group but without recursive investigation of tripartition in its subdivisions.

Simple groups with only two locations are treated as special (elliptical) cases. In these tripartition degrades into bipartition and is represented by the extension of the common edge of the two locations (Figure 4-15c).

Another special case is solitary locations of complex shapes. In these the representation of tripartition is produced by the superimposition of the tripartition lines that are derived under both horizontal and vertical decomposition into horizontal slices (cfr. subsection 2.5.3) from which all lines that fully coincide with full edges of the locations are omitted (Figure 4-15d).

Complex (second level) location groups are investigated with respect to tripartition through the investigation of their parts. Subsequently they fall into two main cases according to whether they are subdivided into non-overlapping subgroups (i.e., simple groups) or not. In complex groups that are subdivided into non-overlapping groups the representation of the tripartition schema is formed by the extensions of the tripartition lines of each subgroup until they meet the perimeter of the whole group. Lines that fully coincide with full edges of the perimeter of the group are omitted because they are unrelated (do not contribute) to the overall correlation of the subgroups, that is, the tripartite organization of the whole group (Figure 4-15e).

In complex groups that are not subdivided into non-overlapping subgroups the representation of the tripartition schema is formed by putting together the tripartition lines of each subgroup as they are within each subgroup, without any extension. Again tripartition lines that fully coincide with full edges of the group perimeter are omitted (Figure 4-15f).

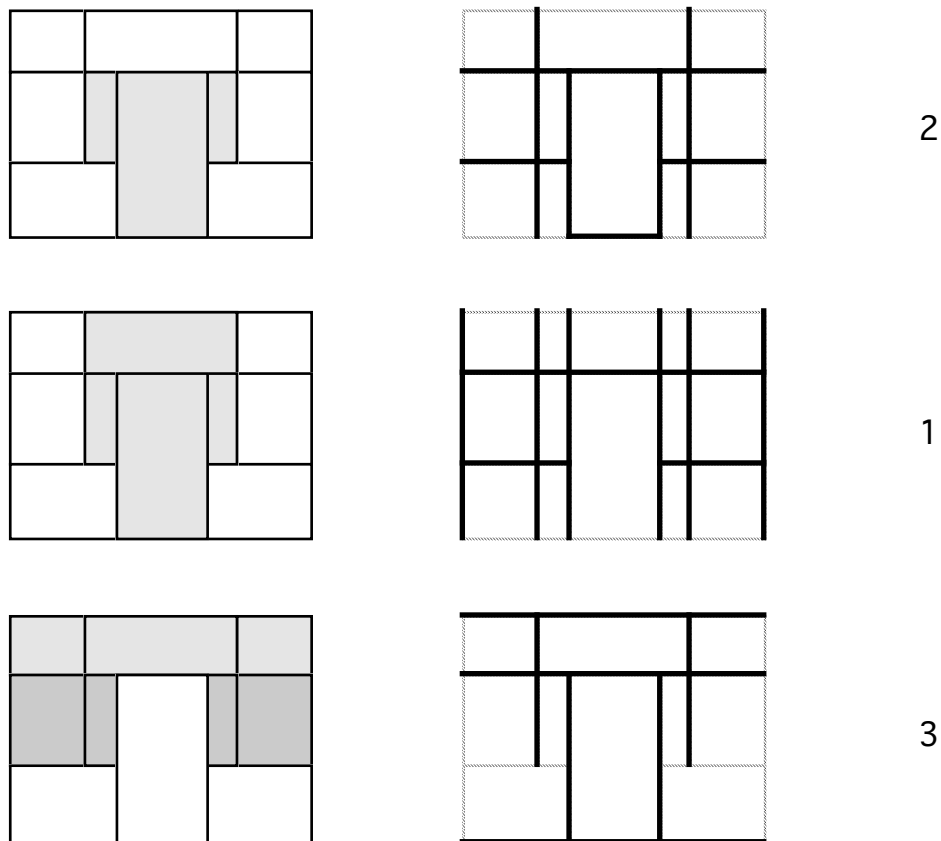
A final particular case concerns groups which include locations of complex shape. In these the the locations of complex shape are reduced into the rectangular slices (vertical or horizontal) that are related to the grouping relationships (Figure 4-21).

#### 4.4.4 Cognitive filtering on the level of the whole architectural plan, I: tripartition

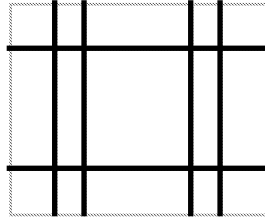
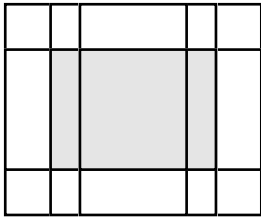
The investigation of tripartition in location groups may be considered as a framework of the evaluation of the well-formedness of these groups with respect to the classical canon and indirectly of the well-formedness of the whole architectural plan. The schema of tripartition is a device of coordination that applies primarily to the whole and, therefore, it is imperative to correlate the local tripartition frames with the global tripartition frame. Investigation of tripartition on the level of the whole plan independently of that on the level of location groups would have been inappropriate because tripartition is not merely a hat to be fitted to a plan by pushing and pulling— any abstract form could in fact be somehow fitted to the dimensionless

schematic architectural plans of chap but this would not justify their existence in a classical plan.

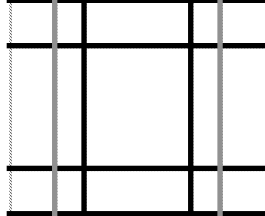
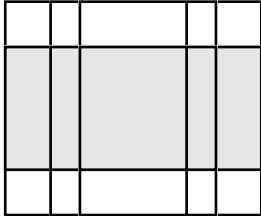
The instance of the tripartition schema in the description of an architectural plan must be verified structurally, by the relations between the tripartition of the parts and of the whole. Such verification may be performed in either of two ways. The first is to identify the tripartition frame of the whole plan and then evaluate its compatibility with the tripartition frames of its location groups. The other way is to correlate the tripartition frames of the groups and evaluate the product of the correlation as the expression of tripartition on the level of the whole plan. Of the two alternatives the latter was preferred because it is easier and faster than the former and at the same time equally effective and reliable, as well as because it conforms to the general bottom-up approach of chap.



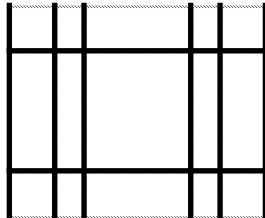
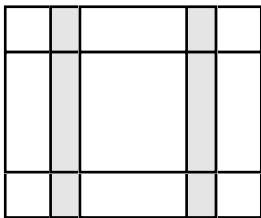
**Figure 4-17** *Tripartition in the case of Villa Zeno*  
*in each row: left: one of the alternative descriptions (cfr. Figure 3-44); middle: the global tripartition frame (black lines denote top level tripartition lines); right: the position of each alternative in the preference order with respect to tripartition (and taxis in general)*



3

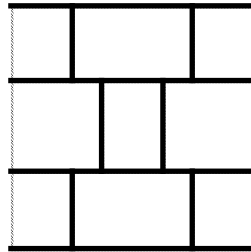
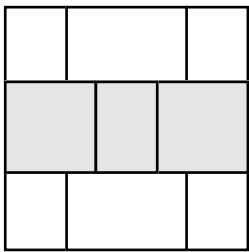


2

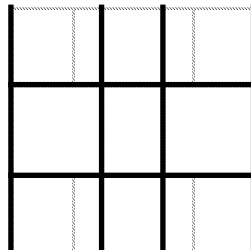
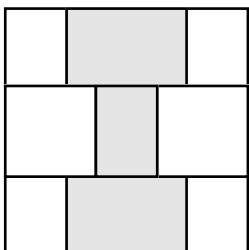


1

**Figure 4-18** Tripartition in the case of Villa Thiene (cfr. Figure 3-45)

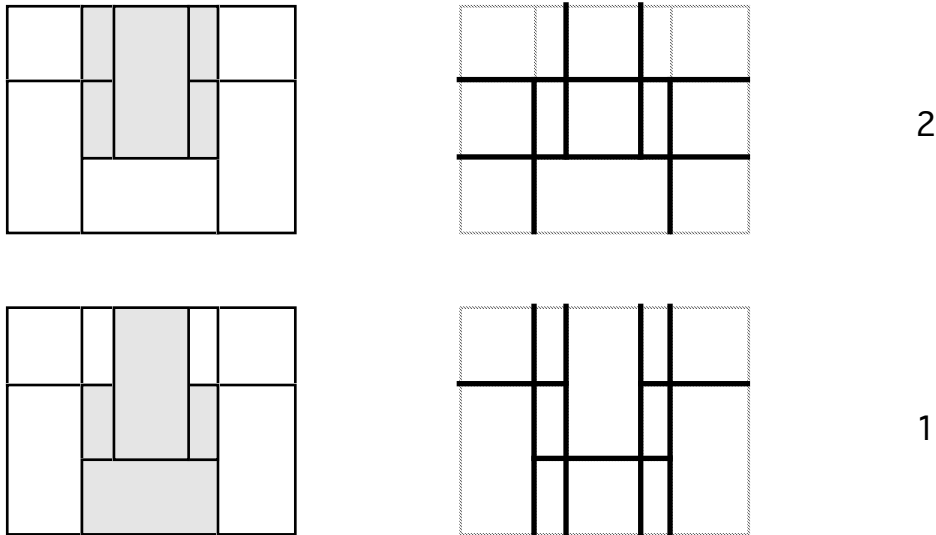


2

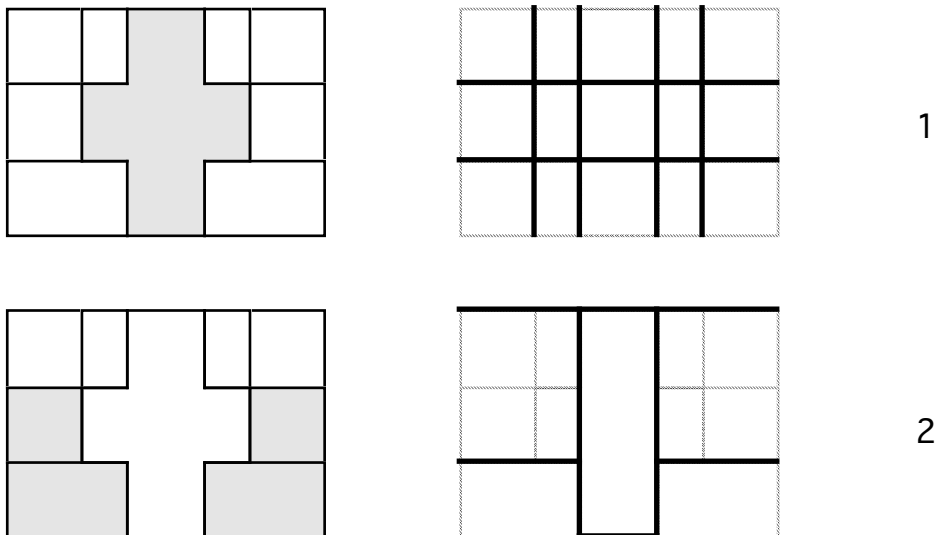


1

**Figure 4-19** Tripartition in the case of Villa Ragona (cfr. Figure 3-46)



**Figure 4-20** *Tripartition in the case of Villa Poiana (cfr. Figure 3-47)*



**Figure 4-21** *Tripartition in the case of Villa Foscari (cfr. Figure 3-48)*

Correlation of local tripartition frames and formation of the global frame is performed by putting together the local frames. The resulting pattern is the representation of the tripartition schema for the particular plan, a derivative of the basic 3 x 3 pattern (Figures 4-17 to 4-21). Evaluation of this global representation of tripartition establishes a preference order for the alternative descriptions of an architectural plan that were derived by the grouping process of Chapter 3.

The evaluation of the global tripartition frame is based on three criteria (in order of significance): (1) proximity to the 3 x 3 pattern, (2) definition of the plan perimeter and (3) type of tripartition lines. Evaluation is performed in the following manner: initially alternative descriptions of the same architectural plan are arranged in a preference order determined by the least number of fusion, addition, deletion or embedding operations required to produce their global

tripartition frame from the 3 x 3 pattern in the manner prescribed by Tzonis & Lefaivre [1986]. The most preferred description is the one that requires the less number of such operations (Figures 4-17, 4-19, 4-20, 4-21). Patterns that require exactly the same number of operations are distinguished by the number of edges of the plan perimeter they include: the greater this number, the higher the preference order of the description (Figure 4-18). Patterns that are still equivalent are distinguished by the number of tripartition lines produced by the tripartition frame of each of the three initial (top-level) subdivisions of a simple location group (or subgroup of a complex plan — Figure 4-15b): the less this number, the higher the preference order of the description (Figure 4-18).

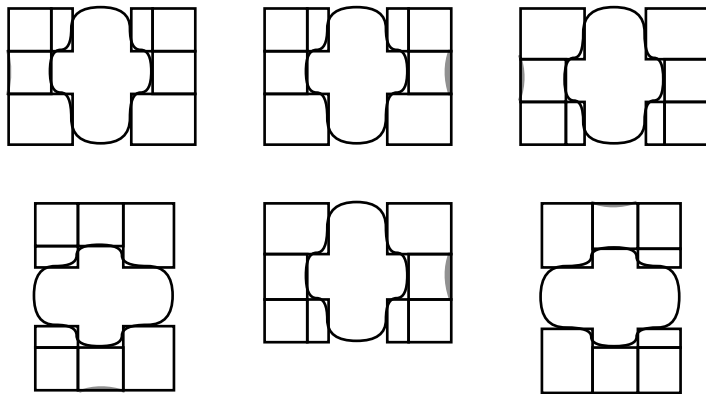
#### 4.4.5 Cognitive filtering on the level of the whole architectural plan, II: symmetry

Subsections 4.4.3 and 4.4.4 considered cognitive filtering with respect to one of the three levels of classical architecture according to Tzonis & Lefaivre [1986], *taxis*. Of the other two, *genera* are beyond the scope and limitations of *chap*. Consequently, the last level, *symmetry*, is also immaterial to *chap* as far as it concerns relationships between elements of the *genera* (building elements). Of the remaining aspects of *symmetry*, the grouping constraints used for the derivation of the description of an architectural plan (Chapter 3) cover most aspects of metric patterns and architectural figures on the level of individual location groups. What remains to be considered is *symmetry* on the level of the whole architectural plan.

In the investigation of *symmetry* in architectural plans *chap* proposes an augmentation of the formulation of the classical canon by Tzonis & Lefaivre by introducing elements used in the investigation of figural goodness through *transformational economy* by Garner [1974] and Palmer [1982, 1983]. In this, the goodness of a pattern is determined by its transformational invariance, that is by its *symmetry* with respect to a particular reference frame. *Symmetry* in the measurement of transformational economy of a pattern is related to its group theoretic structure, that is, to the set of geometric transformations that leave the pattern unchanged [Weyl 1952].

What makes this measure of figural goodness interesting in general and attractive to *chap* in particular is that it is *holistic*, in contrast to the more usual *elementaristic* measures, such as Leeuwenberg's coding

theory, which concentrate on the invariance of features [Palmer 1983, pp.275–276]. This holistic character is well suited to the investigation of symmetry in an architectural plan as a whole, with taxis being related to the reference frame of the transformations. In that respect three alternatives emerge: symmetry of location position, size and shape, symmetry of location group position (independent of group type) and symmetry of position of grouping relationships.



**Figure 4-22 Symmetry**

*the six geometric transformations possible in chap in the case of Villa Foscari;*

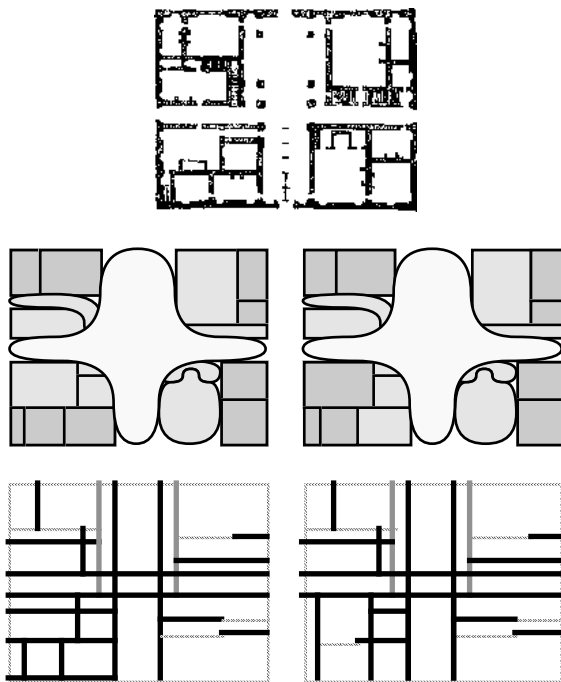
*top left: plan; top middle: reflection with respect to a vertical axis; top right: reflection with respect to a horizontal axis; bottom left: rotation by 270°; bottom middle: rotation by 180°; rotation by 90°*

The first alternative views an architectural plan primarily as a set of locations. In this case symmetry can be measured by the invariance of the plan under the four cases of rotation (including the original position) and the two cases of reflection that are permissible in the totally orthogonal environment of *chap* (Figure 4-22). This type of investigation is certainly relevant to the formal quality of a classical architectural plan but is not pursued in *chap* because it does not relate to the establishment of a preference order for alternative descriptions of an architectural plan as a configuration of location groups. This investigation of symmetry can be used to assign an absolute or relative measure of formal value to different architectural plans but this, as stated previously, is not a concern of *chap*.

The other two alternatives are considered eligible for *chap* because they are based on the description of an architectural plan as a configuration of location groups. Given the precedence of taxis in the formulation of the classical canon by Tzonis & Lefaivre, these two investigations take place with reference to the tripartition schema and its representation, as derived by the process described in the previous subsection. In other words, investigation of symmetry in *chap* follows the investigation of taxis as an elaboration of the relations between the global

tripartition frame and the spatial arrangement of the location groups of an architectural plan. The invariance of the 3 x 3 grid, the embodiment of taxis for  $\text{chap}$ , with respect to the transformations of Figure 4-22 is a clear indication that the global tripartition frame of an architectural plan is the reference frame for the investigation of symmetry. In both the investigation of symmetry of location group position (independent of group type) and symmetry of position of grouping relationships, the position of these entities is considered with respect to the global tripartition frame of the plan.

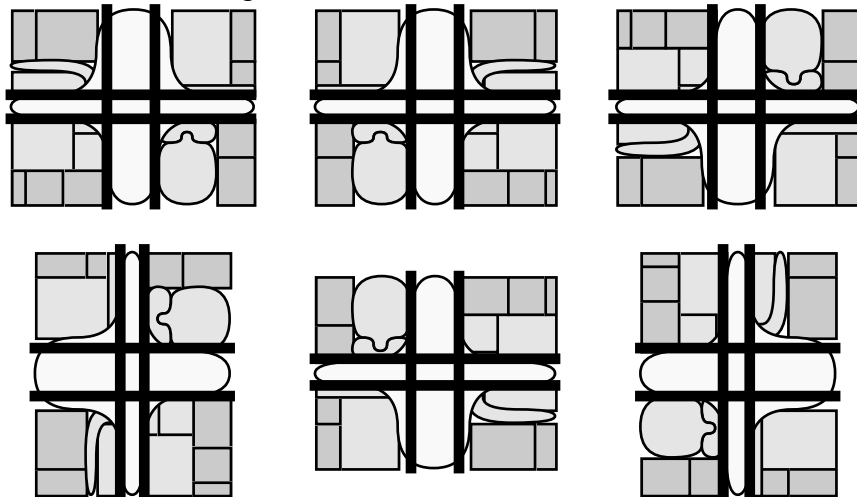
The Palladian villa plans we have used so far as examples of the procedures and processes of  $\text{chap}$  offer poor illustrations for the investigation of symmetry because they are all characterized by symmetry with respect to a central vertical axis, a relationship that also holds for all alternative descriptions derived in Chapter 3 and hence also for the global tripartition frames that correspond to each description. For this reason we shall use another architectural plan to illustrate the investigation of symmetry (Figure 4-23). This plan is only partially symmetrical with respect to a central vertical axis. Its central cross-shaped location is symmetrical to itself. The remaining four areas are also symmetrical in outline but the locations they contain are not. Figure 4-23 depicts two alternative descriptions of the plan and their corresponding global tripartition frames. Let us suppose that these are equivalent with respect to tripartition (an assumption that is very close to the actual evaluation) and proceed to symmetry.



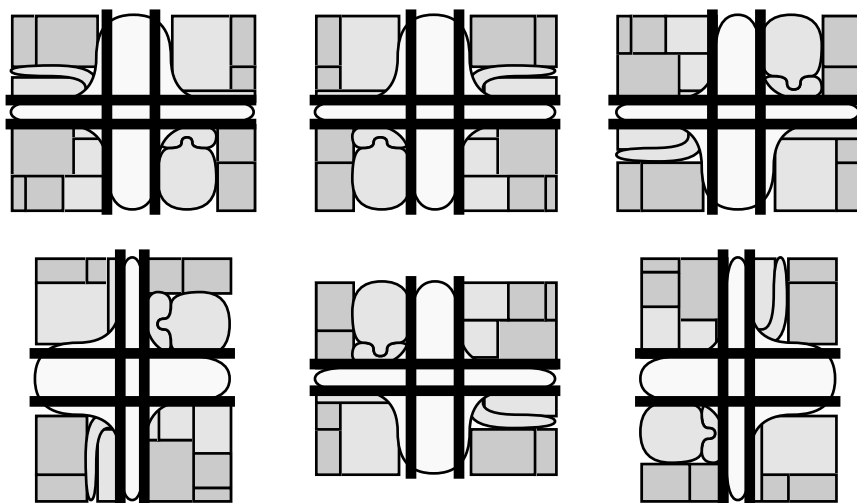
**Figure 4-23** An architectural plan not completely symmetrical with respect to a horizontal or vertical central axis (from Serlio [1619]) from top to bottom: digitized image; two alternative descriptions ('A' (left)

and 'B' (right) — the difference being in the location groups on the bottom left of the plan); corresponding global tripartition frames

In both descriptions the top level element of the global tripartition frames is a clear and complete 3 x 3 pattern whose cells are further subdivided mainly through embedding operations. The transformations of the plan descriptions are considered with respect to this reference frame. In the case of the symmetry of the position of location groups, we can abstract the positions of location groups of the particular plan with respect to which of the halves (top, bottom, left or right) of the cell of the 3 x 3 pattern each group occupies. A comparison of the transformations of the two alternative descriptions on this level (Figures 4-24 and 4-25) reveals that description 'A' is more invariant than description 'B' as its transformations include two pairs of identical arrangements.

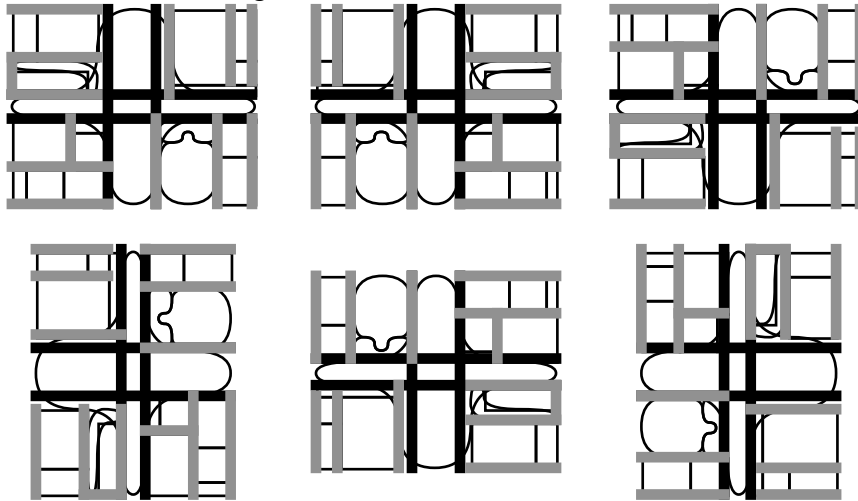


**Figure 4-24** Symmetry transformations of description 'A' of Figure 4-23 with respect to the position of location groups

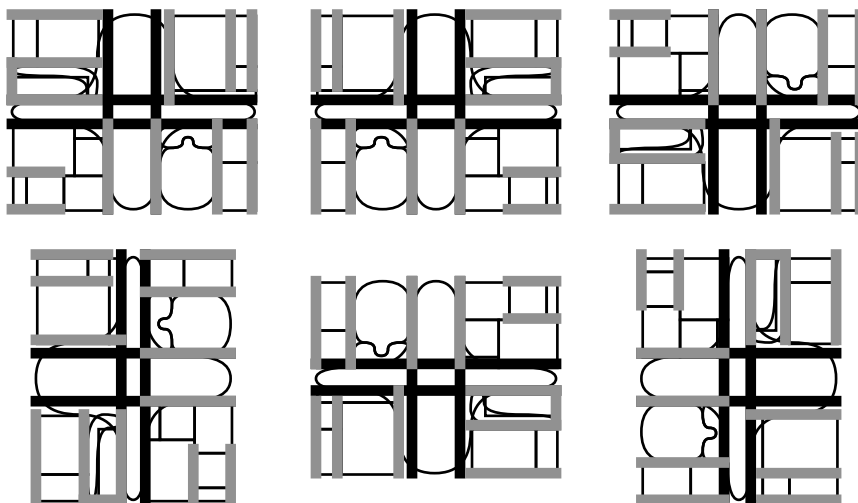


**Figure 4-25** Symmetry transformations of description 'B' of Figure 4-23 with respect to the position of location groups

In the investigation of symmetry of grouping relationships (Figures 4-26 and 4-27) the position of each relationship can be similarly abstracted into a small number of positions relative to the cell a group belongs to by considering a 3 x 3, 3 x 2, 2 x 3 or 2 x 2 grid within each cell. This investigation does not support preference for either description, in all probability due to that all relationships are alignment fronts of short length.



**Figure 4-26** Symmetry transformations of description 'A' of Figure 4-23 with respect to the position of grouping relationships (alignment fronts)



**Figure 4-27** Symmetry transformations of description 'B' of Figure 4-23 with respect to the position of grouping relationships (alignment fronts)

Despite the inconclusive results of the last investigation, symmetry provides chap with a concrete preference for description 'A' which is more clear than that of tripartition in the particular case. As with tripartition, the investigation of symmetry also offers the essential framework for the evaluation of different architectural plans on an absolute or relative scale. Although such evaluations

are no concern of chap, they could also have extensions in the classification of architectural plans: a comparison of plans with a clear 3 x 3 top level element in the global tripartition frame of their preferred description (Figures 4-19, 4-20 and 4-23) suggests that symmetry of the position of location groups and of grouping relationships are no less useful than tripartition of the cells of the 3 x 3 pattern as taxonomic criteria.

## 4.6 Cognitive filtering: conclusions and extensions

The present chapter has introduced a technique for the recognition of the overall type of architectural plans. This technique, cognitive filtering, is unconventional in that it derives the attributes of the type in each instance in a bottom-up manner and because this bottom-up process is a direct investigation of the well-formedness of the instance with respect to its formative context.

Recognition of posture within Marr's computer vision framework illustrates that, however unconventional cognitive filtering may be, it simply amounts to a reversal of the conventional categorization by matching a derived description to a stored model. This reversal is accomplished through the analysis of both the derived descriptions and the models with respect to what essentially is the underlying structure of the model base. Such analysis is beneficial for the variability and flexibility of recognition as it allows the combinatorial definition of the models (or types or categories) that concern recognition, similarly to the way descriptions are formed in most computer vision systems.

For chap the combinatorial definition of types of architectural plans, although not a common device in architecture, conforms to the proscriptive character of the classical canon as suggested by Tzonis & Lefaivre [1986] and accounts for the implicit distinction between local formative operations (such as proportionalities) and global coordinating structures (such as the 3 x 3 and 5 x 3 patterns). This in a sense refutes the conventional view of the classical canon as a generative system (as implied in, for instance, shape grammars). Aspects of the classical canon are selectively involved in all levels of recognition (and hence also of architectural synthesis) but only as determinants of its acceptability with respect to some abstract formal principles and not as guides and modifiers of the functional and constructional requirements that are practically always apparent in local design decisions. This role of the classical canon is evident in the increase of its significance as we

move from the partial to the whole and from the specific to the abstract: on a local level the simplicity and uniformity of spatial arrangement supported by the classical schemata often fades in the face of more pressing requirements (as in the apparent disorder of the four corner areas in the plan of Figure 4-23), while on a regional level and even more on the level of the whole plan the expressions of the schemata are unambiguously present even if subtle or partially disrupted.

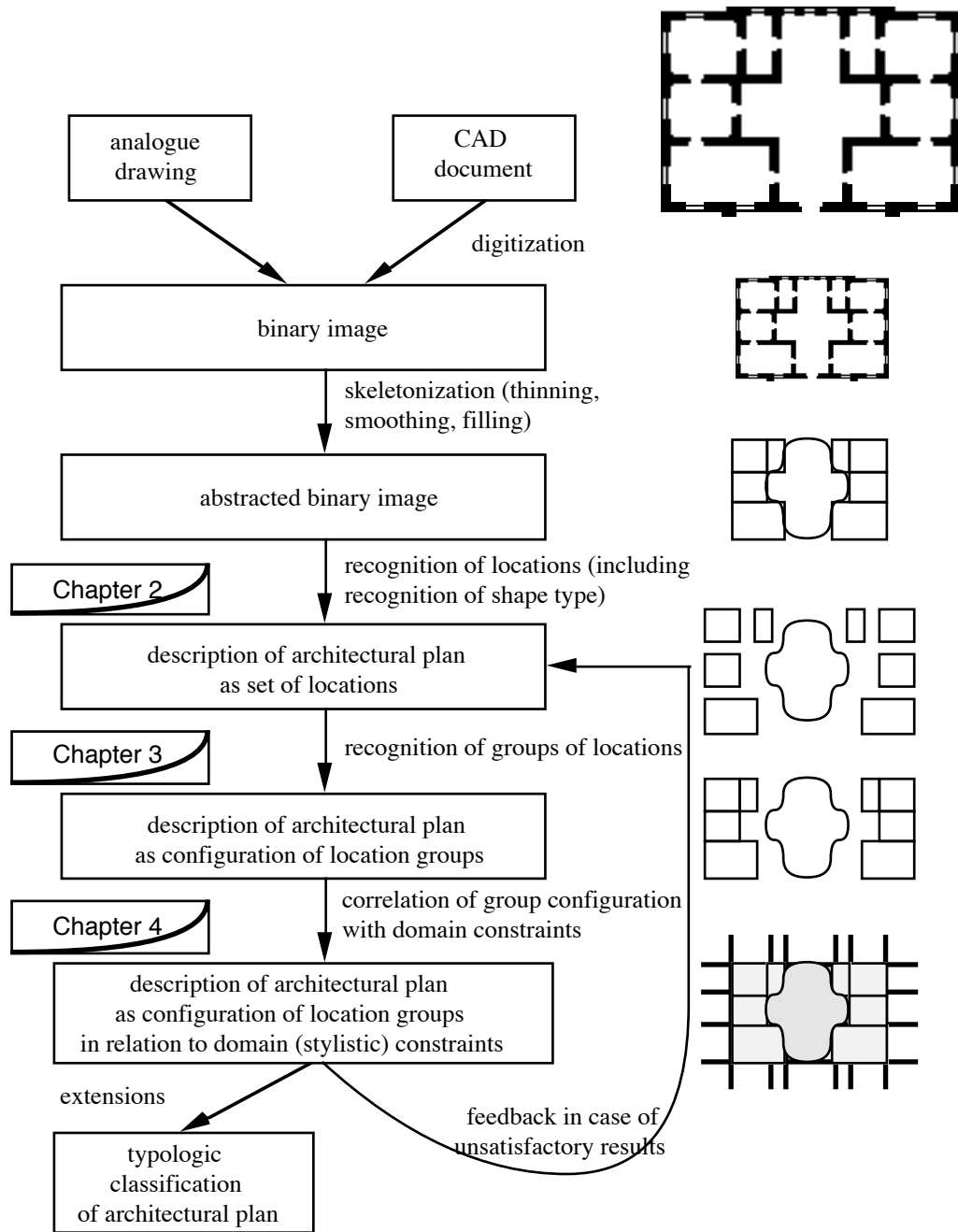
The ability of cognitive filtering to accommodate the classical canon in such a manner relates to its general ability to differentiate between the descriptive and the interpretation levels of recognition. As a result, cognitive filtering can account for the differences in perceived (“substantive”) information in structurally similar images. Black [1972, p.108] observes that measurements of the information content of a picture may be the same or similar in two paintings of a flock of grazing sheep, while “the displayed objects of two such paintings might be manifestly different”. Cognitive filtering is an attempt to distinguish between these two distinct aspects of an image. The configuration of generalized cones in the image of a human figure is often quantitatively the same regardless of the posture and the actions of the figure. This configuration is sufficient for the recognition of the image as one of a human figure but this should not be taken as an indication of the insignificance of the postural aspects of the figure. As we all know from everyday experience, posture supplies far more clues about the intentions of a person than the mere fact that it is a human being.

Similarly, in the context of CHQP, the configuration of location groups could be considered as sufficient information on an architectural plan on the basis of which one could select it from a data base of stored precedents; in fact one could argue that the mere geometric and topologic attributes of each of the locations of the plan, without any grouping whatsoever, would suffice for its selection with respect to, for instance, an adjacency diagram laid out in the manner of space allocation methods or of the dual graph representation. While this is certainly true, it is also ultimately insignificant. Much of our understanding and appreciation of an architectural plan depends on its evaluation from a variety of points of view, among which the formal perspective is often predominant. Omission or even implicit inclusion of formal analyses and evaluations reduce the effectiveness and reliability of the selection process and reduce it into mere search which, however sophisticated in its control and matching components, can be very disappointing in a realistically large collection of precedent architectural plans. The inclusion of extensive formal analyses in the stored descriptions of the plans can transform the collection of stored precedents into a highly flexible and transformable

hierarchy of formal categories which are investigated with respect to their prima facie ability to contain, for instance, some specified adjacency patterns or even the abstract types of spatial articulation they correspond to. This approach is potentially more efficient and also conforms to the proscriptive character of the classical canon or of any other formal rule system in architecture.

A related issue is the nature of the evaluation returned by cognitive filtering. In its present version cognitive filtering in CHAP is designed specifically for the relative evaluation of alternative descriptions of the same plan with the purpose of establishing a preference order for these descriptions. A possibly fruitful extension of this analysis of well-formedness is the development of an evaluation system for different architectural plans on a relative or absolute scale. This is of little concern for recognition but of some significance for the retrieval of a plan from a collection of stored precedents. The cognitive filtering system presented in section 4.4 provides the essential framework for such evaluations, although the correlation of the investigations of taxis and symmetry require extensive further consideration.

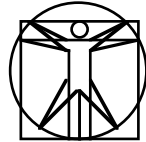
Another extension of the proposed system could be the complete integration of cognitive filtering and thus of recognition in the derivation of a description. The bottom-up structure of cognitive filtering is specifically designed with this integration in mind but in the present versions (cfr. sections 4.3 and 4.4) cognitive filtering is simply superimposed to a derived description. The correspondence between the structure of the description and that of cognitive filtering facilitates recognition but is far from a complete integration.



**Figure 5-1** A schematic outline of chap

# CHAPTER 5

## AN EPILOGUE



### 5.1 Synopsis

Chapter 4 concluded the development of the proposed recognition of architectural plans in the framework of chap, a system for the input and storage of architectural plans in machine environment. Each of Chapters 2 to 4 has described one of the three basic modules of chap. These modules are

- the transformation of the digitized image of an architectural plan into a description of the plan as the set of its locations (Chapter 2)
- the recognition of formal grouping relationships between the locations of an architectural plan and the derivation of one or more alternative configurations of location groups that describe the spatial articulation of the plan (Chapter 3)
- the investigation of the well-formedness of these alternative descriptions and hence of the plan itself with respect to the classical canon (Chapter 4)

The modules are independent from each other in that the procedures of any module do not cooperate directly with procedures of other modules. Each module operates very much on its own, receiving from the preceding module only its final output, that is, a description of the architectural plan as a whole and from all aspects investigated in that module. The second module (the one described in Chapter 3) is invoked only after the first module (described in Chapter 2) has completely analysed the input digitized image of the plan and returned a description of the plan in terms of its locations. The second module uses only this description to arrive at a number of alternative configurations of location groups. Only then is the third module (described in Chapter 4) invoked to analyse the well-formedness these

configurations.

In this serial modular system there is only one feedback possibility (indicated by a feedback link in the diagrammatic outline of Figure 5-1): if the investigation of the well-formedness of the description(s) of the plan in terms of its location groups does not reach a satisfactory result (a case not considered in Chapter 4 because it relates to a general measurement of well-formedness for all plans on a relative or absolute scale), then the recognized locations are re-investigated (return to the first module), in particular within location groups that have caused the unsatisfactory results. Grouping relationships in the new (or reconfirmed) set of locations is then considered once more by the second module. Again emphasis is on the groups that have caused the unsatisfactory evaluation of well-formedness. The resulting (or reconfirmed) configurations of location groups are fed to the third module which reexamines their well-formedness for the final time. If unsatisfactory results are reported again, then either the plan is not a good classical composition (or no classical composition at all) or it indicates a niche not covered by CHAP. A complete record of the results of all procedures of all models facilitates this feedback.

Chapter 1 has indicated that the dissertation is concerned with the development of an approach to the definition and recognition of visual / spatial architectural representations for the computer and not with the presentation of a concrete prototypical computer system. An evaluation of the approach and the techniques proposed in this framework is from many viewpoints premature, as at this stage there can be no reliable testing of CHAP. The only form of evaluation possible at this point is the correspondence of the results of CHAP with intuitive interpretations of the same images, while many of the aspects and elements of CHAP can be justified with respect to human perceptual and cognitive abilities and mechanisms. Both issues relate on the one hand to computer vision, from where CHAP derives most of its techniques, and on the other to the purpose of CHAP and of visual / spatial architectural representations for the computer in general: the development of intelligent computer aids for problem solving in architecture.

## 5.2 The Intelligent Architect, the Architectural Thesaurus and chap

As stated in Chapter 1 (section 1.4), *chap* derives from the general framework of the *Intelligent Architect*, a system for the automated production of design solutions on the basis of a collection of stored precedents, the *Architectural Thesaurus*. *chap* attempts to cover a critical part of the Architectural Thesaurus, the recognition of the formal structure of the architectural plans it contains. This contribution of *chap* to the Thesaurus becomes evident if we consider the differences between the selection of an architectural plan of the Thesaurus and the selection of a plan layout within rectangular arrangements [Mitchell et al 1976; Steadman 1983].

Both the Thesaurus and the exhaustive enumeration of dimensionless plan layouts by rectangular arrangement techniques constitute comprehensive databases of precedents, the difference been in the nature of the precedents. Rectangular arrangements store abstract *types* of plans which could be seen as variations of some generic cases, while the Thesaurus contains actual plans, with a specific performance and a specific potential with respect to the accommodation of each possible set of activities. However, this difference effectively disappears if the description of the plans stored in the Thesaurus is just an adjacency graph which is simply matched to the adjacency graph derived from each set of programmatic requirements, that is, exactly as in rectangular arrangements.

What makes the Thesaurus a more comprehensive tool than rectangular arrangements, shape grammars and similar automated design techniques is integration of high-level formal recognition and analysis, as attempted in *chap*. Recognition of the formal structure of a plan, its spatial articulation and formal subdivisions improves the ability to match programmatic requirements with precedent solutions on more abstract levels and hence more efficiently; it also improves the ability to match partially and by analogy without ignoring formal constraints and interpretations, two issues that are particularly significant for the effectiveness and creativity of an automated design system. *chap* makes explicit the architectural knowledge of form that is contained in an architectural plan, that is, in a representation of a design solution. Integration of this knowledge in the selection of a precedent solution can resolve formal questions directly and thus dismiss the problems observed by Mitchell [1987] with respect to the correlation of formal and functional aspects in systems for the automated production of architectural designs.

## 5.3 Extensions to CAAD

Besides the Intelligent Architect and the Architectural Thesaurus *chop* addresses general aspects of the computerization of architecture, as discussed in Chapter 1. The first of these is CAAD and, more precisely, the debate on the role of computerized drawing in the computerization of architecture. *chop* proposes that, contrary to the general distinction between dumb drawing and intelligent design CAAD inherited from design methods, drawing is an integral part of the design process (at least in the form the latter exists today). In fact, as one can see when approaching architectural drawing from the direction of recognition, it is the primary *representation* used by architects.

*chop* offers several arguments for a precise definition of the role of drawing in CAAD but, more significantly, offers several tools for bringing drawing representation in CAAD closer to the abstract levels that concern not only the users of CAAD systems but also the development of architectural knowledge based systems. Each module of *chop* contains techniques that can be integrated in CAAD systems as a kind of perceptual mechanisms which allow the computer to recognize in a drawing the same formal entities the designer sees and decides about. Explicit representation of the formal consequences of design decisions leads automatically from drawing operations to design decisions, i.e., bridges the gap between the description of the geometry of a design in machine environment and the higher (normative) levels of design thinking.

*chop* indicates the directions for the development of tools for the representation and manipulation of architectural plans as spatial structures, as coherent configurations of locations. This level of representation is absent from existing CAAD tools, despite that practically all automated design approaches in architecture, from space allocation to shape grammars, operate primarily on this level. *chop* suggests a comprehensive technique for the recognition of locations in an architectural plan and techniques for the recognition of their formal relationships and hence offers the means for the integration of such automated design techniques in the transfer of the traditional 'manual' design process to the computer, as well as for the evaluation of automated design techniques and their correspondence to architectural knowledge in that context.

## 5.4 Architectural typology

One of the most interesting recent developments in architectural theory has been a short but inspirational debate

on typology. This debate represented an attempt to go beyond the superficial institutional classifications (house, school, office, hospital) one might encounter in an architectural curriculum or library. Unfortunately the attempts for the definition of new approaches to typology never went beyond abstract considerations on the nature of architectural types and never crossed the border into computer-related architectural research.

The Architectural Thesaurus, as its name denotes, must rely on a consistent and efficient classification of architectural plans from which its internal organization would be derived and according to which queries would be interpreted and search performed. CHAP offers no explicit classification tools but the recognition process of CHAP suggests a very specific approach to the derivation of a typology of classical architectural plans.

Recognition in CHAP is bottom-up. It proceeds from locations, the atomic parts, to the whole plan through grouping on the basis of common spatial relationships. This suggests that a typology of classical plans could be based on a repertory of plan parts, that is, location groups. One could collect a big number of classical plans and record the spatial relationships that group locations together, as well as the formal characteristics of the groups. An appropriate organization of these patterns could result into a more elaborate, detailed and better substantiated version of the grouping constraints employed in Chapter 3. The corresponding repertory of location groups would then allow description of architectural plans on the basis of the type of their parts and their configuration.

However, as Chapter 4 suggests, this is not enough. The configuration of group types must be complemented by a number of appropriate cognitive filters which relate the parts of the plan to the criteria that qualify an architectural plan as a classical composition. These need not be investigated through exhaustive analyses of existing architectural plans but must relate to a general theory about the structure and composition of the classical canon. The cognitive filters represent the essence of the typology in that they explain relations between different types and even the very existence of types in the typology.

CHAP suggests that a typology of architectural plans is *combinatorial* in nature as it consists of types of location groups and types of relationships between these groups which produce a large finite number of plan types. The structure of the typology (especially on the higher levels) relies more on the relationships than on the groups and its content and size are checked by a number of proscriptive constraints on the structure of the groups and on the compatibility of groups and relationships. What makes this approach to the development of architectural typologies of

particular interest to the computerization of architecture is that there are indications that group types remain invariant for several architectural formal systems (if not for all), while relationship types and acceptability constraints are as different as the formal systems they represent. As this approach bears a strong resemblance to the way the characters of an alphabet combine on the basis of a relatively small number of rules to form immensely large numbers of words for a variety of languages, it seems possible that a computational definition of architectural typologies and their input and manipulation in machine environment should follow similar principles and adopt equivalent techniques.

## 5.5 The structure of automated design systems

All research in computer-based systems ultimately amounts to attempts to investigate and define the generative devices that would allow us to automate the production of new design solutions and understand our highly successful intuitive processes. CHAP, although not directly related to automated production of designs, is developed in relation to a specific approach (that of the Intelligent Architect) and therefore is particularly sensitive to this issue.

The position of CHAP with respect to the automated production of design solutions is the same as for architectural typologies: the two phases of recognition, derivation of a description and interpretation with respect to the classical canon, correspond to two distinct modules of an automated design system. The first module consists of the essentially generative operations. These concern the definition of a configuration of instances of specific group types that accommodate the programmatic requirements in activity spaces and their correlations, conform to the various site and construction constraints, and in general satisfy the largest part of the brief.

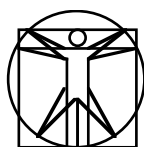
In the history of CAAD, space allocation techniques represent the most extreme overestimation of this module as *the* design process that should concern an architect. Space allocation techniques failed to acknowledge the multiplicity of viewpoints involved in the specification of the functional patterns that should be accommodated in the building and ignored the formal architectural types which define the geometry of the parts of a building. Shape grammars represent the most comprehensive attempt to resolve such problems by constraining the spatial structure of a plan through local formal rules. Still, recognition of formal types that would allow abstraction of the description of the design

solution is only implicit in the local shape rules.

Even more implicit in shape grammars are elements of the second module of an automated design system according to the approach that underlies CHQP. This second module is superimposed on the first and its purpose is to control and coordinate the parts of the design solution that were defined by the first module. The criteria and devices of the second module have nothing to do with the particular programme of requirements. Instead, they relate to abstract formal principles which determine the (aesthetic) quality of the whole composition. In all probability this module operates in a proscriptive manner, in accordance with the nature of the classical canon [Tzonis & Lefaivre 1986, p.6] or any other architectural rule system.

The distinction between two modules, one accommodating the basic, primitive operations which assign clusters of activities to location groups of compatible types, and one accommodating abstract and general formal constraints of coordination and well-formedness, comes into marked contrast with the only automated design technique that addresses explicitly and directly issues of aesthetics, shape grammars. The approach which underlies CHQP does not consider the formal rules of classicism or any other architectural system to be generative in the sense that they guide the design decisions on the spatial articulation of a particular building. On the contrary, it supports that the purely generative process is more or less the same for designing with all formal systems and that it relates more to functional performance, construction and other factors which are only indirectly influenced by formal systems (something that explains, for instance, the transfer of elements of one system to another and the ease of transition of an architect from one system to another). This process is complemented by a number of acceptability constraints which relate more to the design as a whole than to local events. Such constraints supply strong preference criteria for local design decisions (for instance, promote the use of certain location group types) by imposing a coordinating framework which may be subtle but is always present and always effectively guides our perception of a building and of its plan.

# REFERENCES



## A

- J.S. Ackerman, 1977. *Palladio*. rev. ed. Penguin; Harmondsworth [originally published 1967]
- G.J. Agin, 1972. *Representation and description of curved objects*. Ph.D. thesis, Stanford Artificial Intelligence Laboratory Memo AIM-173
- Ö. Akin, 1986. A formalism for problem restructuring and resolution in design. *Planning and Design* **13** 223–232
- Ö. Akin, C.-C. Chen, B. Dave & S. Pithavadian, 1986. A schematic representation of the designers' logic. *CAD and robotics in architecture and construction. Proceedings of the Joint International Conference at Marseilles, 25-27 June 1986*. Kogan Page; London / Nichols; New York
- W. Alonso, 1964. *Location and land use. Toward a general theory of land rent*. Harvard University Press; Cambridge, Massachusetts
- W. Alonso, 1975. Location theory. J. Friedman & W. Alonso (eds) *Regional policy. Readings in theory and applications*. MIT Press; Cambridge, Massachusetts, & London
- C. Alexander, 1964. *Notes on the synthesis of form*. Harvard University Press; Cambridge, Massachusetts
- E. André, G. Bosch, G. Herzog & T. Rist, 1987. Coping with the intrinsic and deictic uses of spatial prepositions. P. Jorrand & V. Sgurev (eds) *Artificial Intelligence II: methodology, systems, applications*. North-Holland; Amsterdam
- F. Arbab, 1987. A paradigm for intelligent CAD. P.J.W. ten Hagen & T. Tomiyama (eds) *Intelligent CAD systems I. Theoretical and methodological aspects*. Springer; Berlin, Heidelberg & New York
- R. Arnheim, 1969. *Visual thinking*. University of California Press; Berkeley & Los Angeles, California
- H. Asada & M. Brady, 1986. The curvature primal sketch. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-8**
- F. Attneave, 1954. Some informational aspects of visual

- perception. *Psychological Review* **61** 183–193
- F. Attneave, 1972. Representation of physical space. A.W. Melton & E.J. Martin (eds) *Coding processes in human memory*. Winston; Washington, D.C.
- F. Attneave, 1982. Prägnanz and soap bubble systems. J. Beck (ed) *Organization and representation in perception*. Erlbaum; Hillsdale, New Jersey
- P.-O. Åstrand & K. Rodahl, 1986. *Textbook of work physiology. Physiological bases of exercise*. 3rd ed. McGraw-Hill; New York

## B

- A. Baddeley, 1986. *Working memory*. Clarendon; Oxford
- M. Balachandran & J.S. Gero, 1988. A model for knowledge-based graphical interface. J.S. Gero & R. Stanton (eds) *Artificial Intelligence developments and applications*. North-Holland; Amsterdam
- D.H. Ballard, 1984. Task frame in robot manipulation. *Proceedings of the National Conference on Artificial Intelligence*. Morgan Kaufman; Los Altos, California
- D.H. Ballard & C.M. Brown, 1982. *Computer vision*. Prentice-Hall; Englewood Cliffs, New Jersey
- H.G. Barrow & J.M. Tenenbaum, 1978. Recovering intrinsic scene characteristics from images. A.R. Hanson & E.M. Riseman (eds) *Computer vision systems*. Academic Press; New York
- J. Beck, 1982a. Introduction. J. Beck (ed) *Organization and representation in perception*. Erlbaum; Hillsdale, New Jersey
- J. Beck, 1982b. Textural segmentation. J. Beck (ed) *Organization and representation in perception*. Erlbaum; Hillsdale, New Jersey
- I. Biederman, 1987. Matching image edges to object memory. *Proceedings 1st international conference on computer vision*. IEEE Computer Society; London
- A. Bijl, 1982. Dumb drawing systems and knowledge engineering. A. Pipes (ed) *CAD 82. Fifth International Conference and Exhibition on Computers in Design Engineering*. Butterworths; Guildford, Surrey
- A. Bijl, 1986a. AI in architectural CAD. *CAD and robotics in architecture and construction. Proceedings of the Joint International Conference at Marseilles, 25-27 June 1986*. Kogan Page; London / Nichols; New York.
- A. Bijl, 1986b. Designing with words and pictures in a logic modelling environment. A. Pipes (ed) *Computer-aided Architectural Design Futures. International Conference on Computer-aided Architectural Design*. Butterworths; London
- A. Bijl, 1987. Strategies for CAD. P.J.W. ten Hagen & T.

- Tomiyama (eds) *Intelligent CAD systems I. Theoretical and methodological aspects*. Springer; Berlin, Heidelberg & New York
- T.O. Binford, 1971. Visual perception by a computer. *IEEE Conference on Systems and Controls*. Miami, Florida
- M. Black, 1972. How do pictures represent? E.H. Gombrich, J. Hochberg & M. Black *Art, perception, and reality*. Johns Hopkins University Press; Baltimore, Maryland, & London
- H. Blum, 1967. A transformation for extracting new descriptions of shape. W. Wathen-Dunn (ed) *Symposium on models for perception of speech and visual form*. MIT Press; Cambridge, Massachusetts
- T.G.R. Bower, 1977. *The perceptual world of the child*. Fontana / Open Books; London
- T.G.R. Bower, 1979. *Human development*. Freeman; San Francisco
- M. Brady, 1982. Parts description and acquisition using vision. A. Rosenfeld (ed) *Robot vision. Proceedings of SPIE – the International Society for optical engineering. Volume 336. March 6–7, 1982, Arlington, Virginia*. SPIE; Bellingham, Washington.
- M. Brady, 1983. Criteria for representations of shape. J. Beck, B. Hope & A. Rosenfeld (eds) *Human and machine vision*. Academic Press; New York & London
- M. Brady & H. Asada, 1984. Smoothed local symmetries and their implementation. *International Journal of Robotics Research* **3** 36–61
- R.A. Brooks, 1981. Symbolic reasoning among 3-D models and 2-D images. *Artificial Intelligence* **17** 285–348
- R.A. Brooks, 1986. Model-based 3-D interpretation of 2-D images. A.P. Pentland (ed) *From pixels to predicates: recent advances in computational and robotic vision*. Ablex; Norwood, New Jersey
- S. Brown, 1989. Opening up design education. *Design Studies* **10** (1) 4–11
- F.E. Brown & J.P. Steadman, 1986. A computerised database of contemporary house plans. *Planning and Design* **18** 405–414
- H. Buffart, E. Leeuwenberg & F. Restle, 1981. Coding theory of visual pattern completion. *Journal of Experimental Psychology: Human Perception and Performance* **7** 241–274

## C

- S. Carlsöö, 1972. *How man moves. Kinesionlogic studies and methods*. Heinemann; London
- M. Clowes, 1971. On seeing things. *Artificial Intelligence* **2** 79–116
- L. Combes, 1976. Packing rectangles into rectangular

- arrangements. *Planning and Design* **3** 3-32
- S. Coren, 1981. The interaction between eye movements and visual illusions. D.F. Fisher, R.A. Monty & J.W. Senders (eds) *Eye movements: cognition and visual perception*. Erlbaum; Hillsdale, New Jersey
- P.P. Cortopassi & T.C. Rearick, 1988. A computationally efficient algorithm for shape decomposition. *Proceedings 1988 IEEE Computer Society conference on computer vision and pattern recognition*. Computer Society Press; Washington, D.C.
- R.D. Coyne, M.A. Rosenman, A.D. Radford, M. Balachandran & J.S. Gero, 1990. *Knowledge-based design systems*. Addison-Wesley; New York
- J.W. Croon, 1986. Thoughts of the architectural thesaurus. Internal report, Faculty of Architecture, Delft University of Technology

## D

- E. Davis, 1983. The MERCATOR representation of spatial knowledge. A. Bundy (ed) *IJCAI-83. Proceedings of the eighth international conference on Artificial Intelligence*. Kaufmann; Los Altos, California
- L.S. Davis, 1986. Two-dimensional shape representation. T.Y. Young & K.-S.Fu (eds) *Handbook of pattern recognition and image processing*. Academic Press; Orlando, Florida
- D.P. Dobkin, 1988. Computational geometry — then and now. R.A. Earnshaw (ed) *Theoretical foundations of computer graphics and CAD*. Springer; Berlin, Heidelberg, & New York
- R.O. Duda & P.E. Hart, 1973. *Pattern classification and scene analysis*. Wiley; New York

## E

- C.F. Earl, 1977. A note on the generation of rectangular dissections. *Environment and Planning B* **4** 241—246
- C.F. Earl, 1980. Rectangular shapes. *Environment and Planning B* **7** 311–342
- R.A. Earnshaw, 1988. Geometric data structures for computer graphics: an overview. R.A. Earnshaw (ed) *Theoretical foundations of computer graphics and CAD*. Springer; Berlin, Heidelberg, & New York
- C.M. Eastman (ed), 1975. *Spatial synthesis in computer-aided building design*. Applied Science; London
- A.S. Espenschade & H.M. Eckert, 1980. *Motor development*. 2nd ed. Merrill; Columbus, Ohio
- G.J. Ettinger, 1988. Large hierarchical object recognition using libraries of parametrized model sub-parts. *Proceedings 1988 IEEE Computer Society Conference*

*on Computer Vision and Pattern Recognition.*  
Computer Society Press; Washington, D.C.

## F

- W.H. Fawcett, 1986. Design knowledge in architectural CAD. *Computer-aided Design* **18** 83–90
- M.A. Fischler, 1978. On the representation of natural scenes. A.R. Hanson & E.M. Riseman (eds) *Computer vision systems*. Academic Press; New York
- M.M. Fleck, 1986. Local rotational symmetries. *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society Press; Washington, D.C.
- U. Flemming, 1986. On the representation and generation of loosely packed arrangements of rectangles *Environment and Planning B* **13** 189–205
- U. Flemming, 1987. More than the sum of parts. *Planning and Design* **14** 323–350
- R. Freedman, 1987. A computer recreation of Palladian villa plans. *Architecture* **17** 58–66
- H. Freeman, 1961. On the encoding of arbitrary geometric configurations. *IRE Transactions of Electronic Computers* June 1961 260–268
- H. Freeman, 1974. Computer processing of line drawing images. *Computer Surveys* **6** 57–97
- R.W. Frick, 1988. Issues of representation and limited capacity in the visuospatial sketchpad. *British Journal of Psychology*. **79** 289–308
- K.-s. Fu, 1974. *Syntactical Methods in Pattern Recognition*. Academic Press; New York
- K.-s. Fu, 1981. Recent progress in syntactic pattern recognition. L.N. Kanal & A. Rosenfeld (eds) *Progress in pattern recognition — volume 1*. North-Holland; Amsterdam
- K.-s. Fu, 1982. *Syntactical Pattern Recognition and Applications*. Prentice-Hall; Englewood Cliffs, New Jersey
- K.-s. Fu, 1986. Syntactic pattern recognition. T.Y. Young, K.-s. Fu (eds) *Handbook of Pattern Recognition and Image Processing*. Academic Press; Orlando, Florida

## G

- W.R. Garner, 1974. *The processing of information and structure*. Erlbaum; Potomac, Maryland.
- W.R. Garner & D.E. Clement, 1963. Goodness of pattern and pattern uncertainty. *Journal of verbal learning and verbal behavior*. **2** 446–452

- P.C. Gasson, 1983. *Geometry of spatial forms. Analysis, synthesis, concept formulation and space vision for CAD*. Ellis Horwood; Chichester
- J.S. Gero (ed), 1977. *Computer applications in architecture*. Applied Science; Barking, Essex
- J.S. Gero, 1986. An overview of knowledge engineering and its relevance to CAAD. A. Pipes (ed) *Computer-aided Architectural Design Futures. International Conference on Computer-aided Architectural Design*. Butterworths; London
- J.S. Gero & M. Maher, 1987. Future Roles of Knowledge-based Systems in the Design Process. T. Maver & H. Wagter (eds) *CAAD Futures '87. Proceedings of the Second International Conference on Computer-aided Architectural Design Futures*. Elsevier; Amsterdam
- J.S. Gero, A.D. Radford, R. Coyne & V.T. Akiner, 1985. Knowledge-based computer-aided architectural design. J.S. Gero (ed) *Proceedings of the IFIP WG 5.2. Working Conference on Knowledge Engineering in Computer-aided Design*. North-Holland; Amsterdam
- J.J. Gibson, 1950. *The perception of the visual world*. Houghton Mifflin; Boston, Massachusetts
- J.J. Gibson, 1966. *The senses considered as perceptual systems*. George Allen & Unwin; London
- J.J. Gibson, 1979. *The ecological approach to visual perception*. Houghton Mifflin; Boston, Massachusetts
- J. Gips, 1975. *Shape grammars and their uses*. Birkhäuser; Basel & Stuttgart
- E. Goldmeier, 1972. Similarity in visually perceived forms. *Psychological Issues* **8** 1–135
- E.H. Gombrich, 1972. The mask and the face: the perception of physiognomic likeness. E.H. Gombrich, J. Hochberg & M. Black *Art, perception, and reality*. Johns Hopkins University Press; Baltimore, Maryland, & London
- R.C. Gonzalez & M.G. Thomason, 1978. *Syntactic pattern recognition. An introduction*. Addison-Wesley; Reading, Massachusetts
- R.T. Green & M.C. Curtis, 1966. Information theory and figure perception: The metaphor that failed. *Acta Psychologica* **25** 12–36
- R.L. Gregory, 1970. *The intelligent eye*. McGraw-Hill; New York
- M. Griffiths, 1974. *Introduction to human physiology*. Macmillan; New York / Collier Macmillan; London
- F.C.A. Groen & R.J. van Munster, 1984. Topology Based Analysis of Schematic Diagrams. *Proceedings 7th Conference on pattern recognition*. July 30 – August 2, 1984. Montreal.
- F.C.A. Groen & R.J. van Munster, 1986. Computer aided analysis of schematic diagrams. E.S. Gelsema & L. N.

Kanal (eds) *Pattern recognition in practice II; International workshop, Amsterdam, June 1985*. North-Holland; Amsterdam.

- A. Guzmán, 1968. *Computer resolution of three-dimensional objects in a visual scene*. Ph.D. thesis, MAC-TR-59, Project MAC, MIT. Cambridge, Massachusetts

## H

- P. Haggett, 1966. *Locational analysis in human geography*. St.Martin's Press; New York
- R.N. Hall, 1983. The use of GABLE OMS (Object Modelling System) in the building design process. W.P. de Wilde, M. Mollaert and H. Bruelinckx (eds) *European computer aided architectural design education (ECAADE). Proceedings of the international conference*. Brussels
- A.R. Hanson & E.M. Riseman, 1978. VISIONS: A computer system for interpreting scenes. A.R. Hanson & E.M. Riseman (eds) *Computer vision systems*. Academic Press; New York
- J.G. Hay & J.G. Reid, 1988. *Anatomy, mechanics, and human motion*. 2nd edition. Prentice-Hall; Englewood Cliffs, New Jersey
- D.J. Heeger & A.P. Pentland, 1986. Seeing structure through chaos. *Proceedings Workshop on motion: representation and analysis*. IEEE Computer Society Press; Washington, D.C.
- M. Herman & T. Kanade, 1986. Mosaic scene understanding system. A.P. Pentland (ed) *From pixels to predicates: recent advances in computational and robotic vision*. Ablex; Norwood, New Jersey
- A. Herskovits, 1970. *On boundary detection*. Memo 183, Project MAC, MIT. Cambridge, Massachusetts
- A. Herskovits, 1986. *Language and spatial cognition. An interdisciplinary study of the prepositions in English*. Cambridge University Press; Cambridge
- T. Higuchi, 1983. *The visual and spatial structure of landscapes*. MIT Press; Cambridge, Massachusetts, & London
- J. Hochberg, 1972. The representation of things and people. E.H. Gombrich, J. Hochberg & M. Black *Art, perception and reality*. Johns Hopkins University Press; Baltimore, Maryland, & London
- J.E. Hochberg & E. McAlister, 1953. A quantitative approach to figural "goodness". *Journal of Experimental Psychology* **46** 361–364
- J. Hofer-Alfeis, 1987. Automated conversion of existing mechanical engineering drawings to CAD data structures: state of the art. K. Bø, L. Estensen, P.

- Falster & E.A. Warman (eds) *Computer applications in production and engineering*. CAPE'86. North-Holland; Amsterdam
- D.D. Hoffman & W.A. Richards, 1986. Parts of recognition. A.P. Pentland (ed) *From pixels to predicates: recent advances in computational and robotic vision*. Ablex; Norwood, New Jersey
- B.K.P. Horn, 1986. *Robot vision*. MIT Press — McGraw-Hill; Cambridge, Massachusetts, & London
- E.M. Hoskins & M.F. Bott, 1972. DHSS: Development planning system and production documentation system. *Proceedings of the International Conference on Computers in Architecture*. British Computer Society; London
- D. Howard, 1980. Four centuries of literature on Palladio. *Society of Architectural Historians Journal* (October 1980) 224–241
- D. Howard & M. Longair, 1982. Harmonic proportion and Palladio's Quattro Libri. *Society of Architectural Historians Journal* (May 1982) 116–143
- D. Huffman, 1971. Impossible objects as nonsense sentences. B. Meltzer & D. Michie (eds) *Machine Intelligence 6*. Edinburgh University Press; Edinburgh

|

- W. Isard, 1956. *Location and space-economy. A general theory relating to industrial location, market areas, land use, trade, and urban structure*. MIT Press; Cambridge, Massachusetts

J

- D. Jain & M.L. Maher, 1988. Combining expert systems and CAD techniques. J.S. Gero & R. Stanton (eds) *Artificial Intelligence developments and applications*. North-Holland; Amsterdam
- C. Jenkins & S. Tanimoto, 1983. Quad-trees, oct-trees and k-trees: A generalized approach to recursive subdivision of Euclidean space. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-5** 533–539
- H. Jeong & B.R. Musicus, 1989. Extracting masks from optical images of VLSI circuits. J.L.C. Sanz (Ed) *Advances in machine vision*. Springer; New York, Berlin & Heidelberg
- F.H. Jones, 1986. *Computer aided architecture and design*. William Kaufman; Los Altos, California
- J.C. Jones, 1970. *Design methods: seeds of human futures*. Wiley; New York
- S.H. Joseph, 1989. Processing of engineering drawings for automatic input to CAD. *Pattern Recognition* **22** 1–11

## K

- G. Kanizsa, 1955. Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia* **49** 7–30
- G. Kanizsa, 1974. Contours without gradients or cognitive contours? *Italian Journal of Psychology* **1** 93–112
- G. Kanizsa & W. Gerbino, 1982. Amodal perception: seeing or thinking? in J. Beck (ed) *Organization and representation in perception*. Erlbaum; Hillsdale, New Jersey
- H.S. Kim, K.H.Park & M.Kim, 1987. Shape decomposition by colinearity. *Pattern Recognition Letters*. **6** 335–340
- J. Kittler, 1986. Feature selection and extraction. T.Y. Young & K.-S. Fu (eds) *Handbook of pattern recognition and image processing*. Academic Press; Orlando, Florida
- W. Klein, 1982. Local deixis in route directions. R.J. Jarvella & W. Klein (eds) *Speech, place, and action: studies in deixis and related topics*. Wiley; Chichester
- W. Klein, 1983. Deixis and spatial orientation in route directions. H. Pick & L. Acredolo (eds) *Spatial orientation: theory, research, and application*. Plenum; New York
- J.J. Koenderink & A.J. van Doorn, 1976. The singularities of visual mapping. *Biological Cybernetics* **24** 51–59
- K. Koffka, 1935. *Principles of Gestalt psychology*. Harcourt Brace; New York
- W. Köhler, W., 1929. *Gestalt Psychology*. Liveright; New York
- J.F. Koegel, 1987. A theoretical model for intelligent CAD. P.J.W. ten Hagen & T. Tomiyama (eds) *Intelligent CAD systems I. Theoretical and methodological aspects*. Springer; Berlin, Heidelberg & New York
- B. Kuipers, 1983. Modeling human knowledge of routes: partial knowledge and individual variation. *Proceedings of the national conference on Artificial Intelligence*. William Kaufman; Los Altos, California
- B.J. Kuipers & T.S. Levitt, 1988. Navigation and mapping in large-scale space. *AI Magazine* **9** 25–43
- R. Krishnamurti, 1986. Modelling Design Decisions. *CAD and robotics in architecture and construction. Proceedings of the Joint International Conference at Marseilles, 25-27 June 1986*. Kogan Page; London / Nichols; New York

## L

- J. Lansdown, 1987. Computers and Visualization of Design Ideas: Possibilities and Promises. T. Maver & H. Wagter (eds) *CAAD Futures '87. Proceedings of the Second International Conference on Computer-aided*

- Architectural Design Futures*. Elsevier; Amsterdam
- B.R. Lawson & J.P. Riley, 1982. ISAAC. A technique for the automatic interpretation of spaces from drawn building floor plans. A. Pipes (ed) *CAD 82. 5th International Conference and Exhibition on Computers in Design Engineering*. Butterworths; Guildford, Surrey
- E.L.J. Leeuwenberg, 1971. A perceptual coding language for visual and auditory patterns. *American Journal of Psychology* **84** 307–350
- E.L.J. Leeuwenberg, 1978. Quantification of certain visual pattern properties: salience, transparency, similarity. E.L.J. Leeuwenberg & H.F.J.M. Buffart (eds) *Formal theories of visual perception*. Wiley; Chichester
- E.L.J. Leeuwenberg, 1982. Metric aspects of patterns and structural information theory. J. Beck (ed) *Organization and representation in perception*. Erlbaum; Hillsdale, New Jersey
- M. Leyton, 1986. Constraint-theorems on the prototypification of shape. *Proceedings AAAI-86. Volume 1. Science*. Morgan Kaufman; Los Altos, California
- R.S. Liggett, 1985. Optimal spatial arrangement as a quadratic assignment problem. J.S.Gero (ed) *Design Optimization*. Academic Press; Orlando, Florida, & London
- R.M. Loughed, D. McCubbrey & R. Jain, 1989. Applying iconic processing in machine vision. J.L.C. Sanz (Ed) *Advances in machine vision*. Springer; New York, Berlin & Heidelberg
- D.G. Lowe & T.O. Binford, 1983. Perceptual organization as a basis for visual recognition. *Proceedings of the national conference on Artificial Intelligence*. William Kaufman; Los Altos, California
- K. Lynch, 1960. *The image of the city*. MIT Press; Cambridge, Massachusetts
- K. Lynch, 1971. *Site planning*. 2nd ed. MIT Press; Cambridge, Massachusetts

## M

- C.A. Mackenzie, 1988. Inducing relational grammars from design interpretations. J.S. Gero & R. Stanton (eds) *Artificial intelligence developments and applications*. North-Holland; Amsterdam
- A.K. Mackworth, 1973. Interpreting pictures of polyhedral scenes. *Artificial Intelligence* **4** 121–137
- M. Magee & M. Nathan, 1985. A rule based system for pattern recognition that exploits topological constraints. *Proceedings 1985 IEEE Computer Society conference on computer vision and pattern*

- recognition*. IEEE Computer Society Press; Silver Spring, Maryland
- L. March, 1976a. A Boolean description of a class of built forms. L. March (ed) *The architecture of form*. Cambridge University Press; Cambridge
- L. March, 1976b. Editorial. *Planning and Design* **3** 1–2
- D. Marr, 1982. *Vision. A computational investigation into the human representation and processing of visual information*. W.H. Freeman; San Francisco
- D. Marr, 1978. Representing visual information — a computational approach. A.R. Hanson & E.M. Riseman (eds) *Computer vision systems*. Academic Press; New York
- D. Marr & K. Nishihara, 1977. Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings Royal Society of London* **B 200** 269–294
- L. Matin, 1982. Visual localization and eye movements. W.A. Wagenaar, A.H. Wertheim & H.W. Leibowitz (eds) *Symposium on the study of motion perception*. Plenum; New York
- T. Maver, 1986. Chairman's introduction to Part Three: Artificial Intelligence and knowledge engineering. in A. Pipes (ed) *Computer-aided Architectural Design Futures. International Conference on Computer-aided Architectural Design*. Butterworths; London
- T. Maver, 1988. Software tools for the technical evaluation of design alternatives. T. Maver & H. Wagter (eds) *CAAD Futures '87. Proceedings of the Second International Conference on Computer-aided Architectural Design Futures*. Elsevier; Amsterdam
- D.M. McKeown, Jr. & J.F. Pane, 1985. Alignment and connection of fragmented linear features in aerial imagery. *Proceedings 1985 IEEE Computer Society conference on computer vision and pattern recognition*. IEEE Computer Society Press; Silver Spring, Maryland
- M.A. Meager, 1972. Computer aids in hospital building. *Proceedings of the International Conference on Computers in Architecture*. British Computer Society; London
- M.A. Meager, 1973. The application of computer aids to hospital building. J. Vlietstra & R.F. Wielinga (eds) *Computer-aided design*. North-Holland; Amsterdam
- G. Medioni, 1982. Segmentation of images into regions using edge information. *Proceedings of the National Conference on Artificial Intelligence*
- G.A. Miller, 1968. The magical number seven, plus or minus two: some limits on our capacity for processing information. *The Psychology of communication. Seven essays*. Allen Lane The Penguin Press; London

- G.A. Miller & P.N. Johnson-Laird, 1976. *Language and perception*. Cambridge University Press; Cambridge
- M.A. Milne, 1982. Solar5, a user-friendly computer-aided energy conserving design tools. A. Pipes (ed) *CAD 82. Fifth International Conference and Exhibition on Computers in Design Engineering*. Butterworths; Guildford, Surrey
- M. Minsky, 1975. A framework for representing knowledge. P.H. Winston (ed) *The psychology of computer vision*. McGraw-Hill; New York
- M. Minsky, 1988. *The society of mind*. Simon & Schuster; New York
- J.R. Mitchell & A.D. Radford, 1987. EAVE, A generative expert system for detailing. *Planning and Design* **14** 281–292
- W.J. Mitchell, 1977. *Computer-aided architectural design*. Petrocelli / Charter; New York
- W.J. Mitchell, 1987. Reasoning about form and function. Y.E. Kalay (ed) *Computability of design*. Wiley; New York
- W.J. Mitchell, 1990. *The logic of architecture*. MIT Press; Cambridge, Massachusetts, & London
- W.J. Mitchell, R.S. Liggett & T. Kvan, 1987. *The art of computer graphics programming. A structured introduction for architects and designers*. Van Nostrand Reinhold; New York
- W.J. Mitchell, J.P. Steadman & R.S. Liggett, 1976. Synthesis and optimization of small rectangular floor plans. *Planning and Design* **3** 37–70
- P.H. Mowforth, J.Jelinek & Z.P.Jin, 1987. An appropriate representation for early vision. *Pattern Recognition Letters*. **5** 175–182
- M.J.M. Muller, 1986. Human texture discrimination. *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society Press; Washington, D.C.
- E. Muybridge, 1899. *Animals in motion*. Chapman & Hall; London
- E. Muybridge, 1901. *The human figure in motion*. Chapman & Hall; London

## N

- V.S. Nalwa, 1988. Line-drawing interpretation: a mathematical framework. *Proceedings 1988 IEEE Computer Society conference on computer vision and pattern recognition*. Computer Society Press; Washington, D.C.
- J.H. Nash, 1982. Graphic Interaction with Database Systems. A. Pipes (ed) *CAD 82. Fifth International Conference and Exhibition on Computers in Design*

*Engineering*. Butterworths; Guildford, Surrey

- H. Neuckermans, 1987. The intelligent pencil: a framework for caad in education. *Design Methods and Theories* **21** 700–716
- R. Nevatia, 1982. *Machine perception*. Prentice-Hall; Englewood Cliffs, New Jersey
- R. Nevatia, 1986. Image segmentation. T.Y. Young & K-s. Fu (eds) *Handbook of pattern recognition and image processing*. Academic Press; Orlando, Florida
- R. Nevatia & T.O. Binford, 1977. Description and recognition of curved objects. *Artificial Intelligence* **8** 77–98

## O

- A.L. O'Connell & E.B. Gardner, 1972. *Understanding the scientific bases of human movement*. Williams & Wilkins; Baltimore, Maryland
- T. Oksala, 1987. Logical models for rule-based CAAD. T. Maver & H. Wagter (eds) *CAAD Futures '87. Proceedings of the Second International Conference on Computer-aided Architectural Design Futures*. Elsevier; Amsterdam
- G.C. Osbourn, 1988. A new approach to machine-based perception of monocular images. *Proceedings 1988 IEEE Computer Society conference on computer vision and pattern recognition*. Computer Society Press; Washington, D.C.
- R. Oxman & J.S. Gero, 1987. Using an expert system for design diagnosis and design synthesis. *Expert Systems* **4** 4–15

## P

- S.E. Palmer, 1982. Symmetry, transformation, and the structure of perceptual systems. J. Beck (ed) *Organization and representation in perception*. Erlbaum; Hillsdale, New Jersey
- S.E. Palmer, 1983. The psychology of perceptual organization: a transformational approach. J. Beck, B. Hope & A. Rosenfeld (eds) *Human and machine vision*. Academic Press; New York & London.
- T. Pavlidis, 1977. *Structural pattern recognition*. Springer; Berlin, Heidelberg, & New York
- T. Pavlidis, 1982. *Algorithms for graphics and image processing*. Computer Science Press; Rockville, Maryland
- T. Pavlidis, 1988. Image analysis and graphics. R.A. Earnshaw (ed) *Theoretical foundations of computer graphics and CAD*. Springer; Berlin, Heidelberg, & New York
- A.P. Pentland, 1986. Part models. *Proceedings IEEE*

*Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society Press; Washington, D.C.

- A.P. Pentland, 1986. Parts: structured descriptions of shape. *Proceedings AAAI-86. Volume 1. Science*. Morgan Kaufman; Los Altos, California
- D.N. Perkins, 1982. The perceiver as organizer and geometer. J. Beck (ed) *Organization and representation in perception*. Erlbaum; Hillsdale, New Jersey
- D.N. Perkins, 1983. Why the human perceiver is a bad machine. J. Beck, B. Hope & A. Rosenfeld (eds) *Human and machine vision*. Academic Press; New York & London
- H.L. Pick, 1984. Cognition and action in development: a tutorial discussion. W. Prinz & A.F. Sanders (eds) *Cognition and motor processes*. Springer; Berlin & New York
- S. Port, 1987. Computer-drafting — State of the Art. T. Maver & H. Wagter (eds) *CAAD Futures '87. Proceedings of the Second International Conference on Computer-aided Architectural Design Futures*. Elsevier; Amsterdam
- N.L. Prak, 1977. *The visual perception of the built environment*. Delft University Press; Delft
- W. Prinz & A.F. Sanders, 1984. Preface. W. Prinz & A.F. Sanders (eds) *Cognition and motor processes*. Springer; Berlin & New York

## R

- A.D. Radford, J.S. Gero, M.A. Rosenman & M. Balachandran, 1985. Pareto optimization as a computer-aided tool. J.S. Gero (Ed) *Proceedings of the IFIP WG 5.2. Working Conference on Knowledge Engineering in Computer-aided Design. Budapest, 17-19 September 1984*. North-Holland; Amsterdam
- A.D. Radford & J.S. Gero, 1985. Multicriteria optimization in architectural design. J.S. Gero (ed) *Design Optimization*. Academic Press; Orlando, Florida, & London
- A. Radford & G. Stevens, 1987. *CADD made easy. A comprehensive guide for architects and designers*. McGraw-Hill; New York
- G. Retz-Schmidt, 1988. Various views on spatial prepositions. *AI Magazine* 9 95–105
- R. Roaf, 1977. *Posture*. Academic Press; London
- I. Rock, 1983. *The logic of perception*. MIT Press; Cambridge, Massachusetts, & London
- I. Rock, 1984. *Perception*. Scientific American Library; New York

- A. Rosenfeld & A.C. Kak, 1982. *Digital image processing*. 2nd ed. Vol.2. Academic Press; New York
- M.A. Rosenman & J.S. Gero, 1985. A system for integrated optimal design. J.S. Gero (Ed) *Design Optimization*. Academic Press; Orlando, Florida, & London
- C. Rowe, 1947. The mathematics of the ideal villa. C. Rowe, 1976. *The mathematics of the ideal villa and other essays*. MIT Press, Cambridge, Massachusetts, & London [originally published in *Architectural Review*]
- S. Ruffle, 1986. Architectural design exposed: from computer-aided drawing to computer-aided design. *Planning and Design* **13** 385–389

## S

- H. Samet, 1984a. The quadtree and related hierarchical data structures. *ACM Computing Surveys* **16** 187–260
- H. Samet, 1984b. A tutorial on quadtree research. A. Rosenfeld (ed) *Multiresolution image processing and analysis*. Springer; Berlin, Heidelberg, & New York
- H. Samet, 1988. An overview of quadtrees, octrees, and related hierarchical data structures. R.A. Earnshaw (ed) *Theoretical foundations of computer graphics and CAD*. Springer; Berlin, Heidelberg, & New York
- H. Samet & A. Rosenfeld, 1980. Quadtree representation of binary images. *Proceedings 5th International Conference on Pattern Recognition*.
- T.G. Schilling, 1987. The impact of computer-aided drafting on design office management. Y. Gardan (Ed) *Architectural design and cad: a selection of papers from MICAD*. Kogan Page; London / Nichols; New York.
- G. Schmitt, 1987. Expert systems and interactive fractal generators in design and evaluation. T. Maver & H. Wagter (eds) *CAAD Futures '87. Proceedings of the Second International Conference on Computer-aided Architectural Design Futures*. Elsevier; Amsterdam
- G. Schmitt, C.-C. Chen, J.-C. Robert & J. Weeks, 1986. OPS5 in architecture: four test cases. *CAD and Robotics in Architecture and Construction. Proceedings of the Joint International Conference at Marseilles, 25-27 June 1986*. Kogan Page; London / Nichols; New York
- R. Seidl, 1986. Artificial Intelligence research at the Department of Architecture. The Intelligent Architects Workstation. The Thesaurus. Internal report, Faculty of Architecture, Delft University of Technology
- S. Serlio, 1619. *Tutte l' opere d' architettura et prospettiva di Sebastiano Serlio Bolognese*. Venice [first published in 1537 (bk I), 1540 (bk II), 1545 (bk III), 1547 (bk IV) and 1575 (bk V)]

- E. Shaviv, 1986. Layout design problems: systematic approaches. A. Pipes (ed) *Computer-aided architectural design futures. International conference on computer-aided architectural design*. Butterworths; London
- E. Shaviv, 1987. Generative and evaluative CAAD tools for spatial allocation problems. Y.E. Kalay (ed) *Computability of design*. Wiley; New York
- W.L. Shebiliske, 1984. Context effects and efferent factors in perception and cognition. W. Prinz & A.F. Sanders (eds) *Cognition and motor processes*. Springer; Berlin & New York
- E. Sheerer, 1984. Motor theories of cognitive structure: a historical review. W. Prinz & A.F. Sanders (eds) *Cognition and motor processes*. Springer; Berlin & New York
- O. Silvén, I. Virtanen, T. Westman, T. Piironen & M. Pietikäinen, 1989. A design data-based visual inspection system for printed wiring. J.L.C. Sanz (Ed) *Advances in machine vision*. Springer; New York, Berlin & Heidelberg
- P. Steadman, 1976. Graph-theoretic representation of architectural arrangement. L. March (ed) *The architecture of form*. Cambridge University Press; Cambridge.
- J.P. Steadman, 1983. *Architectural morphology. An introduction to the geometry of building plans*. Pion; London.
- P. Steadman, 1987. Architecture — spatial layout. J. Rooney & P. Steadman (eds) *Principles of computer-aided Design*. Pitman; London
- G.L. Steele Jr, 1984. *Common Lisp. The language*. Digital Press; Bedford, Massachusetts
- V. Stefanelli, 1970. *Giorgio Vasari il Giovane, La Città Ideale. Pianta di Chiese (Palazzi e Ville) di Toscana e d' Italia*. Officina; Rome
- K.A. Stevens, 1986. Inferring shape form contours across surfaces. A.P. Pentland (ed) *From pixels to predicates: recent advances in computational and robotic vision*. Ablex; Norwood, New Jersey
- G. Stiny, 1975. *Pictorial and formal aspects of shape and shape grammars*. Birkhäuser; Basel & Stuttgart
- G. Stiny, 1976. Two exercises in formal composition. *Planning and Design* **3** 185–210
- G. Stiny, 1980. Introduction to shape and shape grammars. *Planning and Design* **7** 343–351
- G. Stiny & J. Gips, 1978a. *Algorithmic aesthetics: computer models for criticism and design in the arts*. University of California Press; Berkeley, California
- G. Stiny & J. Gips, 1978b. An evaluation of Palladian plans. *Environment & Planning B* **5** 199–206

- G. Stiny & W.J. Mitchell, 1978a. The Palladian grammar. *Planning and Design* **5** 5–18
- G. Stiny & W.J. Mitchell, 1978b. An evaluation of Palladian plans. *Planning and Design* **5** 199–206
- P.J. Szalapaj & A. Bijl, 1985. Knowing where to draw the line. J.S. Gero (Ed) *Proceedings of the IFIP WG 5.2. Working Conference on Knowledge Engineering in Computer-aided Design. Budapest, 17-19 September 1984*. North-Holland; Amsterdam
- J. Summerson, 1980. *The classical language of architecture*. revised ed. Thames & Hudson; London [originally published 1963]
- G. Sunde, 1987. A CAD system with declarative specification of shape. P.J.W. ten Hagen & T. Tomiyama (eds) *Intelligent CAD systems I. Theoretical and methodological aspects*. Springer; Berlin, Heidelberg & New York

## T

- P. Tabor, 1976a. Analysing route patterns. L. March (ed) *The architecture of form*. Cambridge University Press; Cambridge.
- P. Tabor, 1976b. Analysing communication patterns. L. March (ed) *The architecture of form*. Cambridge University Press; Cambridge.
- L. Talmy, 1983. How language structures space. H. Pick & L. Acredolo (eds) *Spatial orientation: theory, research, and application*. Plenum; New York
- M.G. Thomason & R.C. Gonzalez, 1981. Database representations in hierarchical scene analysis. L.N. Kanal & A. Rosenfeld (eds) *Progress in pattern recognition – volume 1*. North-Holland; Amsterdam
- W.B. Thomson & J.K. Kearney, 1986. Inexact vision. *Proceedings Workshop on motion: representation and analysis*. IEEE Computer Society Press; Washington, D.C.
- J. Toriwaki & S. Yokoi, 1981. Distance transformations and skeletons in digitized pictures with applications. L.N. Kanal & A. Rosenfeld (eds) *Progress in pattern recognition – volume 1*. North-Holland; Amsterdam
- J.T. Tou & R.C. Gonzalez, 1974. *Pattern recognition principles*. Addison-Wesley; Reading, Massachusetts
- J. Tsukumo & K. Asai, 1986. Machine printed Chinese and Japanese character recognition method and experiments for reading Japanese pocket books. *Proceedings IEEE Computer Society conference on computer vision and pattern recognition*. IEEE Computer Society Press; Washington, D.C.
- M. T. Turvey, 1971. Contrasting orientations to the theory of visual information processing. *Psychological Review*

- A. Tzonis & L. Lefaivre, 1986. *Classical architecture. The poetics of order*. MIT Press; Cambridge, Massachusetts, & London
- A. Tzonis, S.G.M.E. van Asseldonk & J.S.A. de Brauwere, 1987. *Frames, plans, representation. Conceptdictaat inleiding programmatische en functionele analyse*. Faculty of Architecture, Delft University of Technology
- A. Tzonis & R. Scherpbier, 1985. [Untitled internal report on the Intelligent Architect Project] Faculty of Architecture, Delft University of Technology

## U

- B.J. Underwood, 1972. Are we overloading memory? in A.W. Melton & E. Martin (eds) *Coding processes in human memory*. V.H. Winston; Washington, D.C.

## W

- H. Wagter, 1988. CAD-techniques in architecture and building design, a realistic (over)view. T. Maver & H. Wagter (eds) *CAAD Futures '87. Proceedings of the Second International Conference on Computer-aided Architectural Design Futures*. Elsevier; Amsterdam
- D. Walters, 1986. Selection and use of image features for segmentation of boundary images. *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition. June 22-26, 1986. Miami Beach, Florida*. IEEE Computer Society Press; Washington, D.C.
- D. Waltz, 1975. Understanding line drawings of scenes with shadows. P.H. Winston (ed) *The psychology of computer vision*. McGraw-Hill; New York
- D.L. Waltz, 1978. A parallel model for low-level vision. A.R. Hanson & E.M. Riseman (eds) *Computer vision systems*. Academic Press; New York
- S. Watanabe, 1985. *Pattern recognition: human and mechanical*. Wiley; New York.
- R.J. Watt, 1987. An outline of the primal sketch in human vision. *Pattern Recognition Letters* **5** 139–150
- J.A. Webb & E. Pervin, 1984. The shape of subjective contours. *Proceedings of the National Conference on Artificial Intelligence*. Morgan Kaufman; Los Altos, California
- M. Wertheimer, 1938. Laws of organization in perceptual forms. W.D. Ellis (ed) *A source book of Gestalt psychology*. Routledge & Kegan Paul; London [condensed version of German original: *Untersuchungen zur Lehre von der Gestalt, II. Psychologische Forschung* **4** 301–350]

- H. Weyl, 1952. *Symmetry*. Princeton University Press; Princeton, New Jersey
- T.E. Weymouth, J.S. Griffith, A.R. Hanson & E.M. Riseman, 1983. Rule based strategies for image interpretation. *Proceedings of the national conference on Artificial Intelligence*. William Kaufman; Los Altos, California
- P.H. Winston, 1975. Learning structural descriptions from examples. P.H. Winston (ed) *The psychology of computer vision*. McGraw-Hill; New York
- P.H. Winston, 1984. *Artificial Intelligence*. 2nd ed. Addison-Wesley; Reading, Massachusetts
- P.H. Winston & K.B.P.Horn, 1984, *LISP*. 2nd ed. Addison-Wesley; Reading, Massachusetts
- A.P. Witkin, 1986. Scale space filtering. A.P. Pentland (ed) *From pixels to predicates: recent advances in computational and robotic vision*. Ablex; Norwood, New Jersey
- A.P. Witkin & J.M. Tenenbaum, 1986. On perceptual organization. A.P. Pentland (ed) *From pixels to predicates: recent advances in computational and robotic vision*. Ablex; Norwood, New Jersey
- R. Wittkower, 1952. *Architectural principles in the age of humanism*. 2nd ed. Tiranti; London

## Υ

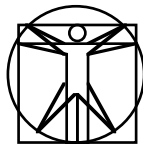
- C.I. Yessios, 1987. The computability of void architectural modeling. Y.E. Kalay (ed) *Computability of design*. Wiley; New York

## Z

- N. Zuech, 1988. *Applying machine vision*. Wiley; New York

# Development of a Computerized Handbook of Architectural Plans

Alexander Koutamanis



## Abstract

The dissertation investigates an approach to the development of visual / spatial computer representations for architectural purposes through the development of the computerized handbook of architectural plans (chap), a knowledge-based computer system capable of recognizing the metric properties of architectural plans. This investigation can be summarized as an introduction of computer vision to the computerization of architectural representations: chap represents an attempt to automate recognition of the most essential among conventional architectural drawings, floor plans. The system accepts as input digitized images of architectural plans and recognizes their spatial primitives (locations) and their spatial articulation on a variety of abstraction levels. The final output of chap is a description of the plan in terms of the grouping formations detected in its spatial articulation. The overall structure of the description is based on an analysis of its conformity to the formal rules of its “stylistic” context (which in the initial version of chap is classical architecture).

Chapter 1 suggests that the poor performance of computerized architectural drawing and design systems is among others evidence of the necessity to computerize visual / spatial architectural representations. A recognition system such as chap offers comprehensive means for the investigation of a methodology for the development and use of such representations.

Chapter 2 describes a fundamental task of chap: recognition of the position and shape of locations, the atomic parts of the description of an architectural plan in chap. This operation represents the final and most significant part of the first stage in processing an image input in machine environment.

Chapter 3 moves to the next significant problem, recognition of the

spatial arrangement of locations in an architectural plan, that is, recognition of grouping relationships that determine the subdivision of a plan into parts. In the absence of systematic and exhaustive typologic studies of classical architecture that would allow us to define a repertory of the location group types possible in classical architectural plans, Chapter 3 follows a bottom-up approach based on grouping relationships derived from elementary architectural knowledge and formalized with assistance from Gestalt theory and its antecedents. The grouping process described in Chapter 3 corresponds both in purpose and in structure to the derivation of a description of an image in computer vision [Marr 1982].

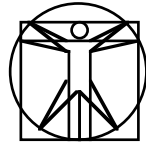
Chapter 4 investigates the well-formedness of the description of a classical architectural plan in an analytical manner: each relevant level (or sublevel) of the classical canon according to Tzonis & Lefaivre [1986] is transformed into a single group of criteria of well-formedness which is investigated independently. The hierarchical structure of the classical canon determines the coordination of these criteria into a sequence of cognitive filters which progressively analyses the correspondence of the descriptions derived as in Chapter 3 to the constraints of the canon.

The methodology and techniques presented in the dissertation are primarily considered with respect to CHAP, a specific recognition system. The resulting specification of CHAP gives a measure of the use of such a system within the context of a computerized collection of architectural precedents and also presents several extensions to other areas of architecture. Although these extensions are not considered as verifiable claims, Chapter 5 describes some of their implications, including on the role of architectural drawing in computerized design systems, on architectural typologies, and on the nature and structure of generative systems in architecture.

## **References**

- D. Marr, 1982. *Vision. A computational investigation into the human representation and processing of visual information*. W.H. Freeman; San Francisco
- A. Tzonis & L. Lefaivre, 1986. *Classical architecture. The poetics of order*. MIT Press; Cambridge, Massachusetts, & London

# Curriculum vitae



## Alexander Koutamanis

### education

1978–85 Aristotle University of Thessaloniki, Greece  
architect ingenieur

### scholarships

1986-89 Greek State Scholarships Foundation

### publications

- 1990 “Architektonische Computer-Vision”  
*Archithese* **20** (3) 44–48
- 1990 “Automatisierung der Architekturausbildung  
durch Computerisierung der Architektur.  
Architectural basics in computerized design  
education at the Faculty of Architecture, Delft  
University of Technology” *Archithese* **20** (3)  
70–72 [with A.Tzonis]
- 1990 “Neue Ansätze im Bereich der  
computerbezogenen Architekturforschung”  
*Archithese* **20** (3) 76–78
- 1989 “Recognition of locations in architectural  
plans” in M.Pietikäinen, J.Röning (eds)  
*Proceedings of the 6th Scandinavian  
Conference on Image Analysis (6SCIA)* Oulu:  
Pattern Recognition Society of Finland
- 1988 The Dictionary of Art (Macmillan) article  
on 19th and 20th century Greek architecture  
[under publication]
- 1987 The Dictionary of Art (Macmillan) articles  
on the architects D. & S.Antonakakis (Atelier  
66), I.Despotopoulos (Jan Despo),  
C.A.Doxiadis, P.Karantinos, L.Kaftantzoglou,  
A.Konstantinidis, N.Mitsakis,  
K.F.Panagiotakos, D.Pikionis, Th.Valentis,

1987 N.Valsamakis, E.Ziller, T.Ch.Zenetos [under  
publication]  
“Perfil de una trayectoria. Vivienda colectiva  
1914–1962” *Monografias de Arquitectura y  
Vivienda* **10** (1987) 46–53 [issue on LeCorbusier  
— with A.Tzonis]

**personal**

1960 born in Thessaloniki, Greece (04.04)  
1984 married to V. Mitossi arch.ing. — one  
daughter (2)

13.08.1990

