Stellingen

behorende bij het proefschrift

Multimedia Tele-education

T. Algra

Delft, 22 februari 1994
1. Moderne tele-educatie technologie zal er mede toe bijdragen dat de traditionele scheiding tussen instituten die afstands-onderwijs verzorgen en die welke behoren tot de categorie kontaktonderwijs langzamerhand zal verdwijnen.

2. Opname van de universiteiten van de Nederlandse Antillen, Aruba en Suriname in een tele-educatie netwerk samen met diverse Nederlandse instituten voor hoger onderwijs zou leiden tot een wederzijdse verrijking van het onderwijspakket en zou tevens kunnen bijdragen tot een evenwichtiger kader-ontwikkeling in de genoemde gebieden.

3. De geringe aandacht in de literatuur voor multimedia-comunicatie via smalbandige kanalen is opmerkelijk, daar te verwachten valt dat voorlopig nog een aanzienlijk aantal potentiële gebruikers op dit type communicatiekanalen zal zijn aangewezen.

4. Differential Chain Coding (DCC) kan beschouwd worden als een sub-optimale variant van Modified Tunstall Coding (MTC).

(dit proefschrift, hoofdstuk 6)

5. In analogie met het fenomeen dat bij beeldtransmissie een minder grote compressie-faktor in sommige gevallen kan leiden tot een kortere overdrachtstijd, dienen opstellers van wetenschappelijke teksten zich af te vragen of de toevoeging van enige redundantie aan de tekst wellicht de benodigde decodeertijd zou kunnen bekorten en derhalve resulteren in een meer efficiënte informatie-overdracht.

(dit proefschrift, hoofdstuk 4)
6. De wanordelijkheid die op veel kantoorbureaus kan worden aangetroffen blijkt uitstekend gehandhaafd te kunnen worden op computerschermen met een gebruikers-interface van het "windows"-type.

7. De toepassing van multimediale documentatiesystemen aan boord van toekomstige bemane ruimtestations leidt tot gewichts- en volumebesparing en verhoogt de efficiëntie van de informatie-overdracht, vergeleken met conventionele op papier gebaseerde systemen, en verdient derhalve serieuze overweging.

8. Kleine gespecialiseerde satellieten ("small-sats") voor aardobservatie-doeleinden kunnen beduidend sneller en goedkoper ontwikkeld worden dan de gangbare complexe satellietsystemen met een veelvoud aan sensoren en functies. Gezien de urgentie van observatie en bewaking van tropische regenwouden is het aan te bevelen dat op korte termijn begonnen wordt met een op dit doel gericht small-sat programma.


9. De gewoonte van de Massoreten in de 6e tot 9e eeuw om bij het kopiëren van de Hebreuwse bijbeltekst tellingen van letters, woorden en verzen in de kantlijn te vermelden, is te beschouwen als één der vroegste toepassingen van systematische foutcontrole ten behoeve van een correcte overdracht van informatie.
MULTIMEDIA TELE-EDUCATION

PC-BASED REAL-TIME NARROWBAND APPLICATIONS
MULTIMEDIA TELE-EDUCATION

PC-BASED REAL-TIME NARROWBAND APPLICATIONS

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
on gezag van de Rector Magnificus,
Prof. Ir. K. F. Wakker,
in het openbaar te verdedigen ten overstaan van
een commissie aangewezen door
het College van Dekanen
op 22 februari 1994, om 14.00 uur

door

Taede Algra

deboren op 10 mei 1952

te Wirdum,

electrotechnisch ingenieur (ir.)
Dit proefschrift is goedgekeurd door de promotor, Prof. dr. J. C. Arnbak

Copyright © 1993 by T. Algra

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Algra, Taede

Multimedia tele-education: PC-based real-time narrowband applications / Taede Algra. - [S.l.:s.n.].-Ill.
Thesis Technische Universiteit Delft - With ref. - With summary in Dutch.
ISBN 90-9006796-5
NUGI 849
Subject headings: distance teaching / telecommunication / computers.

Printed by Bariet, The Netherlands
To my wife Gerdien and my children
Chris, Vanya, Gideon,
Ellis and Joanne
3.2.2 Functional and Performance Requirements 51
3.3 System Design Considerations 65
  3.3.1 Networks and Configurations 65
  3.3.2 Conceptual System Architecture 69
  3.3.3 Critical Technologies 71
3.4 Summary and Conclusions 72
References in Chapter 3 73

4 LOW FRAME-RATE VIDEO CODING FOR NARROWBAND TELE-EDUCATION
4.1 Introduction 77
  4.1.1 Multimedia Tele-education Terminal 77
  4.1.2 Low Frame Rate Video Coding Requirements 79
  4.1.3 Image Compression 80
4.2 Influence of Coder Complexity and Compression Factor on Image Transfer Time 81
4.3 An Effective Algorithm for Low Frame Rate Video Coding (LFVC) 89
  4.3.1 Introduction 89
  4.3.2 Review of Image Compression Methods 89
  4.3.3 Fast and Efficient LFVC Intraframe Compression 102
  4.3.4 Conditional Replenishment 114
4.4 LFVC Codec Design and Performance 120
  4.4.1 Design Considerations and Results 120
  4.4.2 Channel Errors 123
  4.4.3 Concluding Remarks 124
4.5 Summary and Conclusions 126
References in Chapter 4 128

5 FAST AND EFFICIENT VARIABLE-TO-FIXED LENGTH CODING
5.1 Introduction 133
5.2 Tunstall Codes 134
  5.2.1 Introduction to Tunstall Codes 134
  5.2.2 Performance Comparison with Huffman Codes 136
  5.2.3 Software Implementation Aspects 140
5.3 Modified Tunstall Coding (MTC) 145
  5.3.1 The MTC Algorithm 145
5.3.2 Performance 148
5.3.3 Implementation Aspects 149
5.3.4 Sensitivity to Varying Source Statistics 152

5.4 Summary and Conclusions 155
References in Chapter 5 156

6 SOME IMPROVEMENTS OF TELEGRAPHIC CODING APPLIED TO REAL-TIME TELE-EDUCATION

6.1 Introduction 157
6.2 Functional Enhancements 159
   6.2.1 Colour-Fill Facility 159
   6.2.2 Relative Coordinate System (RCS) Facility 162
6.3 Variable-To-Fixed-Length Coding 164
   6.3.1 Relation of DCC to MTC 164
   6.3.2 MTC Applied to Telewriting 167
6.4 Summary and Conclusions 170
References in Chapter 6 171

7 REDUCING END-TO-END DELAYS IN REAL-TIME MULTIMEDIA TELE-EDUCATION APPLICATIONS

7.1 Introduction 173
7.2 End-To-End Delay Reduction Methods 175
7.3 Point-To-Point Pretransfer Protocol 179
   7.3.1 Requirements 179
   7.3.2 The Protocol 180
   7.3.3 Multipoint Pretransfer Protocols 184
7.4 Multicast Pretransfer 184
7.5 Summary and Conclusions 191
References in Chapter 7 193

8 A PC-BASED REAL-TIME MULTIMEDIA TELE-EDUCATION SYSTEM

8.1 Introduction 195
8.2 Objectives, Requirements, and Constraints 196
8.3 Related Developments 198
8.4 System Overview 200
8.5 Multimedia Processing 203
   8.5.1 Page Structure 203
CONTENTS

8.5.2 Text and Graphics 206
8.5.3 Raster-scanned Images 206
8.5.4 Linedrawing 207
8.5.5 Macro Libraries 208
8.5.6 Low Frame-rate Video (LFV) 208

8.6 Communication Services 209
8.6.1 Terrestrial and Satellite-based Configurations 210
8.6.2 Floor Control and Call Set-up 211
8.6.3 Pretransfer Protocols 213
8.6.4 Additional Services 214
8.6.5 Datalink Layer Issues 215
8.6.6 LFV memory occupation 216

8.7 The User's View of the System 218
8.7.1 General 218
8.7.2 Shared-Screen Manipulation 218
8.7.3 Course of the Session 219

8.8 A Classroom Experiment 221
8.9 Summary and Conclusions 223
References in Chapter 8 225

9 SUMMARY AND CONCLUSIONS 227
9.1 Major Results and Conclusions 228
9.2 Recommendations on Future Work 231

LIST OF ACRONYMS 233

APPENDICES

A Software-based implementation of Extended Hammingcode 235
B List of menu command options 237

ACKNOWLEDGEMENTS 247

CURRICULUM VITAE 249

SUMMARY (in Dutch) 251
PREFACE

This dissertation addresses the development of tele-education technology to enable real-time distance courses at a professional quality level. In many countries there are situations in which groups of people lack any form of higher education due to, e.g., geographical barriers. Conventional distance education, in general, is less effective as compared to normal face-to-face courses. Moreover, most regular educational institutes are not adequately equipped to extend their facilities with such a form of distance education. By contrast, real-time tele-education using computer technology in combination with telecommunications, enables a way of teaching approximating the normal face-to-face situation rather closely, and thus can be integrated more easily into the educational organizations of existing institutes. Besides, the high level of interactivity without a long turn-around time contributes to better learning effectiveness.

After surveying tele-education systems and technology, this study describes the design of a tele-education system which is particularly suitable for the extension of the educational facilities of universities to remote communities. The study is focused on the application of third-world small-scale institutes like the University of the Netherlands Antilles (UNA), Curacao, which is used as a case. On the basis of the specific requirements, limiting conditions, and constraints, a PC-based system has been developed, with as much functionality as possible implemented in software. Interactive media considered are low frame-rate video, audio, and the communication of text, graphics, handwriting, and still pictures via a shared screen.

A fast and effective low-frame rate video coding scheme is presented. It achieves acceptable image transfer times for application in real-time tele-education services, where the bandwidths are small to moderate and where a requirement exists for implementation based on standard microcomputer equipment without dedicated hardware. This method includes a novel approach to entropy coding: Modified Tunstall Coding, an improvement of Tunstall's variable-to-fixed-length coding algorithm. It is an attractive choice for software implemented entropy coding in the event of limited alphabet sources with a low entropy.
Further, a number of improvements of telewriting systems is proposed in relation to tele-
education services, concerning a colour-fill facility, an efficient coding algorithm and a
method which supports editing of existing pages. To reduce end-to-end delays of
compound pages for the purpose of tele-education, a method referred to as pretransfer is
developed. It employs periods of low channel utilisation during the session for automatic
transmission of these pages.

The above mentioned techniques and solutions have been incorporated in the design of a
tele-education system appropriate for narrow bandwidths. The development and
implementation of this interactive real-time PC-based multimedia system is described. An
initial version was involved in a classroom experiment at the UNA. From the evaluation
results it may be concluded that the multimedia concept developed meets the requirements
and is a viable approach to tele-education in the specific situations aimed at in this study.
1 INTRODUCTION

1.1 Distance education and face-to-face education

*Face-to-face education* includes all types of education in which teacher and student are present at the same location during the lecture. When this condition is not fulfilled, we speak of *distance education*. For many years already the need for distance education existed: Certain individuals and groups of people could not follow courses at the regular educational institutes because of, e.g., geographical isolation. Well-known examples are countries like Canada, where part of the population is scattered over a large area with difficult travelling in winter time, and Australia, with its extensive areas of low population density. It is not surprising that both countries have a long tradition in the field of distance education.

A more or less generally accepted definition of distance education is (Holmberg [90]¹): "The term distance education covers the various forms of teaching and learning at all levels which are not under the continuous, immediate supervision of tutors present with their students in lecture rooms or on the same premises but which nevertheless benefit from the planning, guidance and tuition (i.e., tutoring, teaching) of the staff of a tutorial organization. Its main characteristic is that it relies on noncontiguous, i.e., mediated, communication. Distance study denotes the activity of the students, distance teaching that of the tutorial organization."

The *correspondence course* is the conventional form of distance education, involving postal delivery of printed educational materials and, in a number of cases, audio- or video-tapes.

¹ In this thesis, the publication year of references will be indicated between brackets
Motivations for students to take a 'distance course', may be the long distance to the educational institute; limited or varying individual possibilities concerning the time available to follow the course; or other personal circumstances (Cerych & Jallade [86], Roberts et al. [89], Algra [88]):

a) **Geographical distance.** Besides the traditional distance-teaching countries Canada and Australia, in many regions in the world a similar situation applies. For instance, the Caribbean Basin, with a multitude of small island communities, lacks higher or even secondary educational institutes in a number of cases. Another example is that of immigrated minorities, which should be able to receive education, given from their home country, in their own language, and in accordance with their own culture. Further, a frequent phenomenon is that of a small target group for a certain course, whose members are scattered over a wide area. This is often the case with courses aimed at continuing specialist education of professionals.

b) **Time constraints.** Another reason to choose distance education may be the fact that one is unable to attend courses by day due to a job or other obligations, leaving only the evening or the weekends. For others, a requirement to complete a course within a certain time is insurmountable. Many forms of distance education do not impose such limitations.

c) **Personal circumstances.** Reasons why some are bound to their home or place of residence may be: personal obligations, disability or financial restrictions (where face-to-face education would require frequent travelling or even a temporary change of dwelling-place).

The drawback of this form of education, as compared to face-to-face education, is the lack of immediate feedback from student to tutor and vice versa. Usually, the course materials are received by mail. After studying a lesson, the student will send back his exercises and questions to the teacher. Answers and corrections will follow one or more weeks later. A long turn-around time has a negative impact on the efficiency of teaching.

Other disadvantages which have been observed are the lack of contacts with fellow students and the feeling of isolation. The issue of pacing students learning at home has been identified as a major problem: it is much more difficult to stimulate, motivate and coach these students (Yule [85]). Usually, the teacher is not informed on the student’s home situation, his background, his reasons to follow the course, his qualities of character and his abilities to complete the offered course level. Mutual interactions between students, common for face-to-face education, are difficult to organise.
1.1 Distance education and face-to-face education

In this context the class interaction process should be distinguished from the individual or sub-group forms of contact. Questions and reactions, resulting from a class or group of students during a lecture, are useful to all the students present. Frequently, questions and issues are discussed which never would have been raised by the students individually. Obviously, this reaction process is also of importance to the teacher. In this way he can complete, clarify, give additional exercises, etc., where necessary. In this process, non-verbal communication (facial expressions, attitude, intonation, etc.) plays a role which should not be underestimated. Contacts between students outside the lecture room are profitable in various ways: mutual assistance in understanding the lessons, cooperation in assignments and exercises, and last but not least the motivating influence of knowing that one has a common goal. This issue has been emphasised by Winders [85] in arguing for the importance of interactive distance education: "The mutual support of knowing that someone else is taking the course is very supportive for home-based students".

Due to these drawbacks, in practice, correspondence courses are suffering relatively high dropout rates. Miller [81] states: "Student isolation is often identified as a factor related to high student failure and withdrawal rates in external study programmes". Meeks [87] mentions that a dropout rate of 70% is not uncommon for correspondence courses at academic level. Only highly motivated students appear to be successful.

Hence face-to-face education is preferable where possible. However, distance education will prove successful and useful in those cases where it is the only feasible way for a student to follow a course, whatever the reason may be.

1.2 Distance education and telecommunication

Recently, the advent of modern telecommunication technologies has changed the prospects of distance education auspiciously. Telecommunication enables real-time links and interaction, eliminating a number of the typical drawbacks of distance education. Advancing technologies based on computer science, micro-electronics and satellite communication stimulated a rapid development of these new educational media during the last decades. When telecommunication plays an essential role in a distance-education system we will use the term tele-education in the following (compare e.g. Berg et al. [92]).

By definition, tele-education encompasses those forms of distance education which employ telecommunication technologies. The primary use of these learning technologies
today has been to replicate the experience of face-to-face instruction. The characteristics of traditional instruction retained are real-time instruction (live), and teacher-student and student-student exchange (interaction). Interaction with the 'teleteacher' is the key ingredient in recreating the traditional instructional model. Whether live or delayed, interaction with the instructor is considered by many as a necessary condition for successful distance education (Roberts, et al. [89]). Other reasons for the implementation of a tele-education service may be: teacher shortage, immediacy (fresh and accurate information), standardisation and convenience (Heinich & Molenda [89]). In this context standardisation means offering of uniform, quality-controlled instruction to learners scattered over many sites or over a large area. Convenience is related to the provision of training courses to learners in service (at a company's premises, etc.).

Barker et al. [89] distinguish correspondence-based distance education (including one-way radio and TV broadcasts) from telecommunication-based distance education (i.e., tele-education). However, in sub-categorizing the latter, their definition appears to be too limited: only two-way real-time audio and/or video systems are included. Here we shall divide tele-education systems into two main categories, for historical and technical reasons. The first category has its roots in conventional distance education: the different modern variants of the correspondence course. The disadvantage of a long response time can be eliminated by electronic means. Since a personal computer is an affordable item for a growing number of people, courses are increasingly offered by electronic mail (E-Mail). Although student and teacher do not interact with each other in real-time, from an educational point of view this way of communication is more effective. Another typical problem of the correspondence course, the lack of contacts with fellow students, can be covered for a great deal by a computer-mediated conferencing facility, a feature enabling students to exchange and display messages by use of a common electronic postbox known as a bulletin board. This category of tele-education systems will be referred to as non real-time or mail-based tele-education.

The second category is that of real-time tele-education and has originated from institutes and universities providing traditional face-to-face education. Now, by use of modern technologies, their educational facilities can be extended over wide areas. Moreover, in this way these institutes are able to support each other by exchanging teachers virtually, offering a broader range of courses than in the past. Applications of this form of tele-education differ from the first category, since lectures and courses are conducted in real-time, using media like video and audio.

Systems and methods can be further sub-categorized according to criteria like interactivity and employed media types. There exist one-way systems without interaction, two-way systems providing various interaction facilities continuously, and systems which
only offer a feedback service on a limited basis. Furthermore interaction may be limited to student-teacher communication or also include mutual student contacts. Media types can be sub-divided into audio and visual media, where the latter comprise motion video, still video (or slow scan television = SSTV), and/or text & graphics. In a two-way situation the system may be symmetrical, meaning that for both directions the same set of media is available with the same performance and functionality, in contrast to an asymmetrical system. Systems may be based on a point-to-point or on a multipoint configuration. The systems functionality and the equipment at the reception sites may service remote groups or individual (e.g., home-based) students. A schematic classification of these different configurations is presented in Fig. 1.1.

In countries like the Netherlands, with its large diversity of educational institutes, distance education will be applied for continuing education (e.g., the Open University), but also in secondary and even primary education. Specialist courses which cannot be taught at every secondary school, could be made widely available by tele-education. Further, Cerich & Jallade [86] identify remedial teaching applications, in order to supplement the regular face-to-face lessons. The vocational education is considered as one of the prime areas for tele-education. The existing educational infra-structure is not able to meet the progressing needs related to the continuously changing technologies. Hence,
tele-education could provide an indispensable instrument for the modernization of the economy and the labour market. Furthermore, in general, a growth of continuing education and open learning is foreseen, of which the needs can be met by tele-education solutions to a large extent.

Obviously, in countries where geographical barriers obstruct the general provision of education, the application of tele-education is even more attractive. This concerns extensive areas with a low population density, or small, mutually isolated, communities. Another possibility is knowledge transfer from universities in developed countries to educational institutes in third-world countries. In this way the level of education in the latter areas could be improved, contributing to the reduction of the technological and economical gap between the developed and the developing countries.

1.3 Scope and Overview of the Thesis

In this thesis we are interested in situations where the application of real-time tele-education can be an appropriate solution. For this purpose the case of the Netherlands Antilles, a small archipelago in the Caribbean Basin, will be analyzed, since it may be considered being representative for a large group of similar situations. These involve small-scale universities in third-world countries, of which the educational facilities should be extended to a limited number of isolated communities. Often, these situations are characterized by typical constraints such as a limited telecommunication infrastructure, and a low level of serviceability with respect to advanced equipment. These small-scale universities, in most cases still in an on-going process of upgrading the quality and diversity of their regular educational programs, usually cannot provide large financial budgets and extensive human resources for the purpose of extra services like tele-education. As will be argued, real-time tele-education systems are the best solution for the application of distance-teaching in these circumstances, since such systems may be employed as an extension of existing face-to-face programs and methods.

The goal of this study is to analyze the requirements which should be met by such a system, taking into account various limiting conditions and constraints. A real-time tele-education system is designed, based on these requirements. The emphasis is on the applications aspects, in the sense that it will be assumed that the system is built on top of standard communications services. Formulated in terms of the ISO-OSI reference model (refer to, e.g., MacKinnon et al. [90]), this study deals with functions corresponding to the uppermost three communication layers, i.e., the application layer, the presentation layer, and the session layer. Lower-layer services will not be discussed, except for some
specific issues which affect the application. Especially, attention will be focused on a number of technical problems due to the application of multimedia concepts in combination with narrow-bandwidth links and largely software-based system functions.

This thesis is organized as follows. Chapter 2 presents an overview of the various tele-education methods implemented in real educational situations, including a number of prominent examples. Regarding the application of novel technologies in education, the importance of a systematic and problem-oriented approach will be stressed. Some general considerations and guidelines are given with respect to the design and implementation of a tele-education solution: various distinguishable tele-education network configurations, a summary of the analysis and design process steps based on the educational technology approach, and a short introduction to telematics and multimedia, resembling the computer-related building blocks of such a system.

In Chapter 3 the University of the Netherlands Antilles (UNA) is presented as a case representing a broader range of similar situations, in order to analyze problem areas and to study various approaches to a solution. It will be argued that real-time tele-education offers the best prospects to realize effective and successful educational services for the benefit of remote communities. Therefore a requirement analysis is performed aimed at the development of a multimedia PC-based system including audio, graphics and reduced video services. Furthermore the impact of the local telecommunication infrastructure on such a system will be assessed and two main network concepts identified: a terrestrial multipoint network and a satellite-based concept. Finally a conceptual system architecture is discussed and a number of critical technologies are identified, to be assessed in the next chapters.

Chapter 4 presents a Low Frame-rate Video Coding method (LFVC), which has been developed as a sub-system for a real-time tele-education terminal. It encompasses a fast and efficient coding scheme, appropriate for software implementation, and aimed at narrow transmission bandwidths. The coding scheme is useful for use in a multimedia tele-education terminal, where there is a requirement for implementation on standard microcomputer equipment without dedicated hardware. Such a subsystem allows the teacher to watch the classroom (and vice versa) during the session, by near full-screen image sequences.

Chapter 5 addresses fast and efficient variable-to-fixed-length coding schemes. High-speed Huffman coding, implemented in software, requires large amounts of memory. It will be shown that Tunstall coding can be applied with the same, or better, performance with respect to the coding rate (depending on the source statistics), but with significant lower complexity. A modified version of the Tunstall coding method is proposed (MTC), with a coding rate closer to the source entropy due to a faster convergence rate. For MTC, complexity aspects are discussed, as well as the sensitivity to
variations of the source statistics. It is shown that MTC is an attractive choice for software-implemented entropy coding in the event of limited alphabet sources with a skew probability distribution.

Chapter 6 proposes a number of practical improvements of present telewriting systems aimed at tele-education applications. It is shown that Differential Chain Coding (DCC), a technique used to encode telewriting, can be enhanced with a fill facility in such a way that graphical editing functions remain feasible. The extension or editing of on-screen-displayed objects by means of an electronic writing tablet implies an ergonomic (eye-to-hand coordination) problem. A Relative Coordinate System (RCS) facility will be proposed by which the user is able to position handwritten graphics on screen at a precisely defined location. Further, it is shown that DCC is a particular, but sub-optimal form of Modified Tunstall Coding. Optimization of DCC by extending the set of codewords results in a significant reduction of the coding rate.

In many situations, transmission of multimedia documents is still constrained by network bandwidth. To reduce end-to-end delays, various approaches may be adopted. Chapter 7 proposes a time-shifting method, referred to as pretransfer, which exploits periods of low channel utilisation during the session for automatic transmission of page sequences. Application protocols are discussed for both point-to-point and multipoint configurations. It is shown that in multicast situations characterised by a large (unnumbered) number of reception stations, pretransfer can be realized by means of appropriate page-scheduling methods.

Chapter 8 describes a PC-based implementation of a real-time multimedia tele-education system in combination with narrowband telecommunication links, suitable for point-to-point and for multipoint configurations. The design meets the requirements set in Chapter 3 and applies the technologies described in the Chapters 4, 5, 6, and 7. A classroom experiment at the university premises at Curaçao is reported, with evaluation results and recommendations.

Finally, in Chapter 9, a summary of this dissertation and some conclusions are given. Recommendations for future work are also presented.

The different chapters are written as more or less independent, self-contained articles. This allows a reader to select subjects of particular interest. For instance the reader who is only interested in educational technology may skip the technologically oriented chapters 4-7. On the other hand, those interested in data compression subjects may prefer to limit themselves to Chapter 4 or 5. The Chapters 6 and 7 give self-contained presentations on telewriting technology and multimedia delay-reduction techniques, respectively.
References in Chapter 1

Algra, T. [88], "Towards a tele-education network on the Netherlands Antilles and Aruba", University of the Netherlands Antilles, UNA-Cahier, No.27, (in Dutch), Curaçao, 1988


Holmberg, B. [90], Perspectives of research on distance education, 2nd edition, Zentrales Institut für Fernstudienforschung, Hagen


2 TELE-EDUCATION: SYSTEMS AND TECHNOLOGY

2.1 Introduction

In Chapter 1, tele-education was defined and the rationale for it was discussed. Table 2.1 gives an outline of the various forms according to which tele-education has been implemented. Note that the table does not list the mode of telecommunication (for instance, a satellite link or a microwave network), since this is purely an implementation issue. Further, it should be taken into account that in practice hybrid systems are often used: in addition to the electronically delivered media, course materials are sent by mail or study meetings are held where the students should be physically present. Real-time tele-education may be used in combination with E-Mail and computer conferencing. The next section presents a survey of the existing methods, and discusses the extent to which the drawbacks of conventional distance education have been eliminated in these applications. Note that it is not the intention to present an exhaustive survey of possible tele-education systems, but to describe the different variants by a selection of systems implemented in real educational situations. Emphasis is put on real-time tele-education, since in the next chapter it will be argued that immediate interactive communication is a mandatory requirement for the situations aimed at.

2.2 A survey of tele-education systems

2.2.1 One-way tele-education

Since the start of radio broadcasting, educational and informative radio programs have been used to reach large target groups scattered over extensive areas. This method enables highly professional broadcasts, by assistance of top experts and dealing with
Table 2.1 Classification of tele-education systems

<table>
<thead>
<tr>
<th>Method</th>
<th>Real-time</th>
<th>Interactive</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-way audio</td>
<td>+</td>
<td>-</td>
<td>Radio</td>
</tr>
<tr>
<td>One-way video</td>
<td>+</td>
<td>-</td>
<td>Television</td>
</tr>
<tr>
<td>Two-way audio</td>
<td>+</td>
<td>+</td>
<td>Radio and/or telephone</td>
</tr>
<tr>
<td>1-way video &amp; 2-way audio</td>
<td>+</td>
<td>+</td>
<td>TV + telephone</td>
</tr>
<tr>
<td>Two-way video</td>
<td>+</td>
<td>+</td>
<td>Television</td>
</tr>
<tr>
<td>SSTV &amp; 2-way audio</td>
<td>+</td>
<td>+</td>
<td>Telephone</td>
</tr>
<tr>
<td>Graphics &amp; 2-way audio</td>
<td>+</td>
<td>+</td>
<td>PC, datalink &amp; telephone</td>
</tr>
<tr>
<td>Electronic class</td>
<td>+</td>
<td>+</td>
<td>PC, datalink</td>
</tr>
<tr>
<td>E-mail course</td>
<td>-</td>
<td>+</td>
<td>PC, datalink</td>
</tr>
<tr>
<td>Virtual class</td>
<td>-</td>
<td>+</td>
<td>PC, datalink</td>
</tr>
</tbody>
</table>

topical subjects. Further, school radio is used to enhance and complement the course materials used in regular face-to-face education. Clearly, this medium is not sufficient to conduct a complete educational program in an efficient way: There is no form of interaction, and no visual component. Actually, this medium does not differ from the use of audio-tapes in conventional distance education, except in the ability to treat highly topical subjects.

Educational television programs have been in use since a few decades. Basically, similarly to one-way audio, this medium does not differ greatly from conventional distance education by video-tape distribution. A video-tape can be viewed at any chosen time. However, a lecture by television gives a feeling of more presence, and does provide the ability to deal with topical subjects. Ellis & Mathis [85] investigated the difference in educational efficiency between live lectures (face-to-face education) and video-taped lectures. Both series of lectures were a part of the course program of Minot State College (North Dakota), and 117 students were involved. They concluded that there was no measurable difference in study results. However, it should be realized that the tapes were produced using the *candid classroom technique*: The interaction process in the classes was also recorded, and benefited the students following the video-taped lectures. Another remark to be made is the possible influence of the novelty effect in an experiment carried out once (Conboy [84], Shimizu & Maesako [88]). A novel experimental form of learning often results in a (short term) positive attitude of the student and influences the research results.

As an illustration we mention here the SITE-project (Satellite Instructional Television Experiment) in India, during the years 1975 and 1976, described as the largest
satellite communication project of all times. NASA provided transponder capacity of the ATS-6 satellite, which covered six of India's states, representing 2330 villages, each with fewer than 3000 inhabitants. Average audiences were 100 per television set or approximately 200,000 nationally. The programs concentrated on upgrading and extending education, health, nutrition, population, agriculture and national integration. The programs were supported by the provision of printed materials, radio broadcasts and personal instructions (Hosie [87]). In the republic of China, the National University Television Network has been in operation for a number of years, reaching 2000 TVRO (Television receive Only) locations with educational and cultural programs (Bouck [87]).

2.2.2 Two-way audio

The first and most obvious applications of tele-education is found in the group of two-way audio systems: The interactive audio contact between students and teacher is realized by means of radio or telephone links. Because of the lack of a visual component, in practice this method is mainly applied as a supplement to correspondence-course materials. At some institutes the method is used in face-to-face education to interview an expert at a distance during the lecture, while the whole class may be involved in the discussion (Heinich et al. [89]).

The Plymouth Audio Conferencing Network (PACNET) at the Plymouth Polytechnic, is described by Winders [85] (Fig. 2.1). The heart of the system is a bridge which enables up to 18 telephone lines to be connected together. Participants ring in to the bridge at an agreed time and are welcomed by the operator and put into the conference.
The operator monitors and supports the group throughout the conference. Loudspeaking telephones enable a group to use one line. A useful application is, e.g., a class questions hour at a regular time in the week as a supplement to a correspondence course. The result is that home-students not only communicate with the teacher, but also with their fellow students. In this way students from all over Great Britain make use of PACNET. An evaluation study reports: "The sessions were seen as useful, fairly productive and meaningful, lively and interesting; and the system was easy to use. Over three-quarters [of the students, T.A.] were satisfied, and two-thirds liked teleconferencing. The third who expressed some dislike of the system may have done so because of technical problems".

In Australia, five universities offer about 850 tele-education courses for some 20,000 students (Guiton [85], Barrett [85]). Where these numbers do not justify the use of media like television, mainly printed materials, audio cassettes and home study kits are available. However, at present these media are supported more and more by telephone contacts between teachers and students. Australia has an established reputation in the field of tele-education. Since 1951 the 'School of the Air' is in operation, enabling children who live in isolation to participate in daily classes with a teacher and fellow pupils by means of two-way short-wave radio.

In 1984 Murdoch University started a project using the FM subcarrier technique, involving the addition of an extra audio signal to an existing FM broadcast carrier (known as SCA = Subsidiary Communication Authorization, see Carlson [75] p. 283). Murdoch University broadcasted the courses as an SCA signal to various groups of students, who were able to follow the programs by use of a dedicated decoder. Telephone feedback by 'phone-in' to the radio studio enabled two-way communication. A second line was available for interruption. Hence all participants were able to listen to the conversation, but direct communication between the various groups is not possible. The objective of these experiments was to investigate the feasibility of this technique. Therefore all sessions were recorded on tape and evaluated later on. A remarkable result was that these radio tutorials increased the amount of student and tutor talk, at the expense of silence, and resulted in longer utterances by students than expected in face-to-face seminars (Barrett [84]).

Dunnett [85] described the DUCT project (Diverse Use of Communications Technology) of the Educational Department of South Australia. The system consists of a terminal with six microphone inputs and an audio amplifier with loudspeaker, which can be connected to a telephone line. Furthermore remote groups are equipped with a response keypad, containing buttons assigned to the individual students. The keypad signals can be interpreted by the teacher observing the signalling lamps of a response indicator. By this equipment it is possible to tutor one or more remote groups of students.
2.2.3 One-way video with audio return

This method involves television broadcasting of lectures via a terrestrial network or by satellite. During the lecture, students may react by phone-in and raise questions. The video signal requires (expensive) broadband broadcast channels. Hence, practical application is only justifiable in the case of large numbers of students. The method is known as Instructional Television (ITV).

The Stanford University system, in service since 1969, is a prime example of ITV. The system provides four video channels from the campus to 120 classrooms at locations within a 35 mile radius. Each channel has a return audio channel so the students can talk directly with the instructor. Students enrolled in these courses on average get good grades; a Stanford University study found that they perform at least as well as full-time students (Baldwin [84b]).

From a growing number of organizations offering ITV courses, we mention the National Technological University (NTU), a partnership of 42 universities, providing both graduate and continuing technical education to working professionals. The courses are delivered by satellite. NTU is fully accredited and offers masters' degrees in eleven different technical disciplines. Most of the graduate and all of the continuing education courses are delivered live and interactively, enabling the remote students to interact via audio with the instructors. This means that the remote classrooms receiving the courses are extensions of the delivering classrooms (candid classroom technique). Additionally, non real-time interaction between faculty and students may be by telephone, E-Mail and facsimile. Currently, there are 375 receive locations throughout the USA, corresponding to 136 organizations receiving the courses. Until January 1992, all satellite transmission was in analog mode. Today NTU delivers all courses in compressed digital mode over 5 channels, with the ability to deliver up to 12 channels per satellite transponder (Baldwin [84a], DeSio [92]).

Other networks are STEP in the North West of the U.S., KNOW in Canada, and the European networks Eurostep and EuroPACE. STEP (Satellite Telecommunications Educational Programming Network) was created to exchange highschool credit courses between institutes in subjects with a lack of certified teaching personnel. Currently STEP operates with over 100 subscribers in 12 states (Roberts, et al. [89]). KNOW (Knowledge Network Of the West communications network) provides distance education since 1980 in British Columbia (Hosie [87]), servicing 250 communities by satellite-delivered programming. About 10% of these programs is reported as interactive. Eurostep is a European association of universities and organizations, employing DBS (Direct Broadcast Satellite) capacity to deliver tele-education. Eurostep uses the enhanced television quality standard
D2MAC (an intermediate between the current PAL standard and the future European HDTV standard) including 4 parallel digital audio channels (Chaplin [89], Champness [90]). EuroPACE is a distribution system for high-level advanced training programs received by company professionals as well as university staff and post-graduate students. Besides telephone, non real-time interaction is possible by a computer conferencing service (Frisk [89]).

2.2.4 Two-way video

Two-way video is the closest imitation of the traditional classroom that present technology allows. The Shar-Ed Video Network, linking four school districts in the Oklahoma Panhandle, is one example. Various courses are exchanged between the schools via fiber-optic leased lines (Roberts, et al. [89]). Another example is described by Shimizu & Maesako [88]. The Tokyo Institute of Technology has two campuses 25 km apart, linked together by optical fiber so that a lecture delivered at one campus could be simultaneously transmitted to the second using four TV channels. The image of the lecture actually being given is taken by two TV cameras and projected life-size on two large video screens in the other lecture room 25 km away, while the image of the students in the latter room is simultaneously transmitted from two TV cameras to two monitor TV sets on the teacher’s desk. This arrangement enables the teacher to observe the performance of students in both rooms while lecturing. The system is designed to be operated by the lecturer on his own. A simple but effective switching configuration allows to change between video sources such as blackboard, overhead projector, slide projector and video-tape recorder, or to change the image on the large screen. Students appreciated the system highly because it enabled them to save commuting time from one campus to another.

Obviously, the visual feedback from students to the teacher is very useful. It allows the teacher to discern how his presentation is comprehended and to repeat or to clarify some issues if necessary. The possibility of non-verbal communication undoubtedly reduces the feeling of distance between teacher and student. However, the high investment and operational costs mostly lead to the choice for a one-way video system.

2.2.5 One-way slow-scan-video (SSTV) and two-way audio

In many situations, budgetary constraints allow only narrow bandwidth communication links, excluding the use of full motion video. However, employing one or more telephone lines allows the transmission of SSTV signals. In this way a still-video image can be
transferred in about 30 seconds, enabling the remote class to observe images of the
teacher, the blackboard, etc. A problem is that the sender should obviously delay talking
about the new image until it is complete and displayed (Bretz [84]).

The University of Wisconsin Extension (Madison) has been described as the world
leader in teleconferencing (Winders [86]). During 1983, 83,000 students enrolled to
follow distance teaching. At 26 locations distant lectures can be attended via audio and
SSTV links. Each site is equipped with three cameras: one to present the teacher, another
to display images of the students, and a third used as a document camera. The teacher
operates the equipment, and the audio signals and images can be exchanged between all
the locations.

Another example is the UWIDITE network (University of the West Indies Distance
Teaching Experiment), described by Lalor [84] and Lalor & Marrett [86]. The campuses
of the University of the West Indies (UWI) are located at Jamaica, Trinidad and Bar-
bados. Besides, a number of smaller islands belong to the service area of the UWI.
UWIDITE is in operation since the mid-Eighties, enabling distance teaching and com-
munication between staff members. Use is made of 4-wire telephone links, interconnec-
ting the islands via satellite, microwave or troposcatter links (see Fig. 2.2). The 'telecon-
ferencing rooms' accommodate audio equipment with at least six microphones and two
loudspeakers, SSTV-equipment, a telewriter, and communication equipment. The applied
SSTV system has three camera inputs and a two-frame memory. However, only one
camera is used, except for the Mona-campus at Jamaica, where an additional camera is
used at a vertical stand as a document camera. Lalor states: "Slow-scan television has
turned out to be the workhorse" and "Students seem to enjoy being introduced to the
tutors and other students by this medium, and it has been much used as a teaching aid
similar to the overhead projector". A telewriter is used as a replacement of the traditional
blackboard. Such a unit consists of a tablet with stylus, a microprocessor and a monitor.
Handwriting is digitized and transmitted to all network locations in real-time and
displayed at the different telewriter monitors. These units have a storage capacity of 15
images, which can be retrieved instantaneously.

Concerning the network operations, much attention has been given to organization,
management and training aspects. Training of teachers is necessary, since the medium
differs significantly from traditional face-to-face education: "While it is inappropriate to
try to design a single method to be used by all teachers in the various situations, the
guiding principle has been that maximum use should be made of the interactive potential
of the system. Teachers were asked to bear this in mind at all times; to try to ensure that
the support materials were ready early; to encourage the students to read and study the
material before each class; and to encourage questions and discussions. The suggestion
was that the teacher should seldom talk for more than 10 - 15 minutes without a break for discussions. The success in achieving these aims varied a great deal. Some teachers have not converted to the interactive style, some have done splendidly. Some classes have not responded well, other have done superbly. The extent of interactions is itself a complex of interactions between subject matter, teacher, class, extent of preparation of students (and teacher), technical quality of the system and no doubt other variables*. UWIDITE is in use for the following applications: inter-campus assistance (some undergraduate courses are given by a teacher residing at one of the other campuses); extra-mural activities (courses, workshops and seminars which are of interest for people at all the islands); training-courses for laboratory instructors; first year B.S. programs in social sciences and law; course for the Certificate of Education; teleconferencing for university staff.

2.2.6 Two-way graphics & text and two-way audio

The British Open University was one of the first institutes realizing two-way audio- graphic tele-education. McConnel [86] states that the Cyclops project has been the first large-scale distance-teaching system in the world to apply the new information technologies. Cyclops is based on conventional television sets, standard audio cassettes, microcomputer technology and telephone. Besides partaking of interactive discussions, Cyclops participants are able to exchange written and hand-drawn information (real-time)
using a lightpen and an electronic writing tablet. The information is written on a shared screen, displayed on all connected TV-sets. These screens are transmitted via telephone links at 1200 Baud, and can be generated at one location and altered, extended, etc., by users on another location. The audio cassette is used as a screen storage medium. In this way the teacher is able to compose screens prior to the session, and retrieve and display those screens at the desired time. McConnel [86] describes an extensive evaluation of this technology, where 16 study centres and 600 students were involved. In general, the lightpen feature was seen as a good possibility to respond in a non-verbal way (e.g., to point to some item on the screen). The electronic writing tablet was much used, and appeared a more handy aid for writing or drawing than the lightpen. A disadvantage observed was the difficulty to coordinate the hand-writing on the tablet with the screen information. The participants concluded that the basis for the success of the system was the use of the TV-screen in addition to audio communication. However, some teachers commented that the screen dimensions were not large enough for the display of handwritten information. Advantages mentioned as compared to other tele-education techniques like ITV are:

1) the relatively simple system concept: no cameras, no studio lightening, no technicians need to stay standby;
2) a graphics facility to summarize discussions and lessons by handwritten information and to present figures and diagrams;
3) the interactive nature of the system: both student and teacher may access the shared screen and transfer information;
4) the system is simple to use and only requires a minimal amount of training;
5) materials can be generated easily as compared to, e.g., the production of video;
6) the operational costs are low.

A second example of a tele-education system using two-way audio in combination with graphics and text, is the system of the University of the Virgin Islands (UVI). Two regular telephone lines are used for audio communication and data, respectively. The data link connects two identical microcomputers in such a way that the operating system of the computer in the remote class is slaved to the teacher's computer. Consequently, all actions and responses on the teacher's computer screen can be followed by the students, including the execution of application programs. This approach proved to be especially useful for computer language instruction. However, a compelling requirement is that all files and software to be activated during the lecture should be transferred to the remote computer prior to the session. The limited capacity of the public telephone network in the
eastern Caribbean area forced to convert to UHF line-of-sight (LOS) radio communication. Thus, the radius of the covered area measures only a few tens of kilometres. The capacity of the applied LOS system equals 24 channels, giving a prospect to extend the communication by a compressed video signal (Rosenthal [87]).

McDonnell [92] describes a network providing real-time distance teaching services for the Confederation of Universities of Central America (CSUCA). Basically, the concept is similar to the UWIDITE network (refer to the previous paragraph) using regular telephone lines. In addition, graphics may be called up on the monitors on all remote locations by a command from the computer of the transmitting centre. These graphics screens should be downloaded to the computers of the remote locations via a computer network, in advance to the scheduled session.

2.2.7 Non real-time tele-education

Various educational institutes have adopted a quite dissimilar approach to distance education. As a basic medium the microcomputer is used, offering telematic services such as computer conferencing and E-Mail. However, the home-based student (in general only individual students are involved) should possess a microcomputer with datacommunication equipment to connect to the institute’s mainframe via one of the public data networks. Teacher-student or mutual student interactions can be realized by the above mentioned services. Systems simulating the traditional class processes, but without real-time connections, have been called virtual class (Hiltz [86]). The advantages over conventional correspondence course methods are: faster transfer of lecture materials, questions and responses. Still most systems lack a form of graphics communication. Further future extensions will involve the transmission of voice mail, audio and video sequences. Of the numerous implementations and trials of these type of systems, we mention two prominent examples.

The New Jersey Institute of Technology’s Electronic Information Exchange System (EIES) is used to offer graduate and undergraduate courses by virtual class. During the course the instructor assigns various projects and assignments that the students are required to complete and comment on. Daily participation in the course is encouraged. To obtain a passing grade, each student is required to input at least two messages per week. Furthermore, students have E-Mail and real-time conferencing capabilities (Meeks [87]).

TeleLearning Systems of San Francisco operates a similar system, although conferencing facilities are not reported: the Electronic University Network (EUN). By the Protege software package, the student is provided with a completely menu-driven and user-friendly communication tool, where essential functions like, e.g., "send this note to
the instructor of this particular course", can be performed by one keystroke. Currently, 200 colleges and universities participate in the programs that form EUN and provide EUN with more than 200 on-line courses (Meeks [87]).

The Dutch Open University experimented with non real-time tele-education courses. According to Bouhuis & Wijnen [88] no significant improvement of learning effectiveness was measured in comparison to correspondence-based distance education.

2.2.8 Electronic class

Computer-mediated distance teaching of a class in real time, has been called electronic class (Centini [86], Scigliano & Centini [85]). NOVA University, Fort Lauderdale, offers various graduate programs to students all over North America, with access to a packet-switched network. In a Unix environment, services are available like E-Mail, computer conferencing, information retrieval (e.g., library information), and computer-assisted instruction packages. Moreover, real-time classes can be taught, involving 26 students at a time. Teacher and students communicate by a text-based split screen: the upper half displays the text input of the teacher, while the lower part is used for the responses of the individual students. Obviously, this way of interaction can be very lively and results in an improved feeling of 'social presence', as compared to the virtual class method. However, course requirements include the obligation to be physically present at some seminar weekends.

2.3 The role of educational technology

The use of advanced technology in education involves serious risks. During the last decades we have been deluged with all kinds of new media: ITV, audio and video tape recorder, telematic services, the microcomputer and the interactive video disk. For each of these media very useful educational applications undoubtedly exist. But in some cases the idea has settled that as soon as the technology (the hardware) has been acquired, the immediate efficient and effective use of it is secured. This is a misunderstanding, which negatively influenced a large number of promising projects, sometimes even with a complete failure as a result.

A striking example is the STD project (Satellite Technology Demonstration), see Hosie [87]. The intention was to prove the value of satellite technology concerning
educational projects for the benefit of the rural population in the western states of the U.S. STD was aborted after one year of experimenting, even before the operational phase was entered. "The nomenclature preferred by writers about the project - Satellite Technology Demonstration - provides insight into the reasons for its failure; the project concentrated on testing the feasibility of operational satellite technology. Concentrating on the technological dimensions of satellites after they had been operating successfully for 10 years was a gross error of judgement" ... "The evaluators noted that software took second priority to hardware and recommended that future projects give media emphasis to program content over pre-testing. The need for instructional support to assist large viewing audiences was also evident. The STD project demonstrated clearly that embarking on a project without a well focused purpose invites disaster".

Various authors warn against this technology-driven approach and plead for a result-oriented and systematic treatment of educational problems. Dunnett [86]: "One unfortunate situation in this process is that the new communication technologies are arriving faster than we are able to manage and develop their usage. This will result in mismanagement if the fact is not realized". Ely & Plomp [86]: "In the past, some of the less successful uses of educational technology occurred when it was offered as a solution to a problem which had not been clearly defined. In the early days media (and to some extent computers today) were viewed as solutions looking for a problem rather than the other way around. The mystique which surrounds the new technologies causes enthusiasts to try to apply them in almost any setting without, however, raising the 'right' questions. It is far better to define and describe the problems facing a country, an organization or an individual and then consider alternative solutions which may involve technology. Unless this view is held by technology's leaders and implementers, most technological innovations are doomed to failure".

Some of the reasons mentioned by the same authors why some projects did not fulfil their expectations are:

a) **Goals were confused.** People participating in the project did not know why educational technology was being used. The hidden objective was to prove the value of a specific medium or technique.

b) **Emphasis on the medium.** Greater importance seemed to be attached to the equipment used than to the design of the program or accompanying materials.

c) **Lack of support systems.** Without a management support system, including logistics on a day-to-day basis, project objectives become difficult to achieve. Furthermore the more sophisticated equipment requires an adequate service organization.

d) **Lack of skills.** Some people, designated to use new technologies do not receive
appropriate training or, at worst, do not receive any training whatsoever.

e) *Expense.* In some cases project implementers are not aware of the expenses that will accrue during the course of operation of the project. If new money is allocated and results are not documented in time, it is likely that financial support for the project will be cut.

f) *Lack of quality software.* After costly and visible equipment is installed, many project participants realize that they must create the 'software' to use with the hardware. Existing software have to be adapted to the new technology or even complete redesign is necessary. If this issue does not get enough attention in time, the best hardware equipment will not lead to the expected educational results.

g) *Lack of system focus.* Some projects are concerned with only one (or a limited number of) aspects of the problem rather than the totality. Attention is given to only a few aspects of a program and not to the entire operation.

The threads which run through successful programs can be summarized as follows: these programs meet *critical* educational needs (such as shortage of teachers or escalating costs); they are *cost-effective*; the delivery systems are relatively simple and available; the emphasis is on the *design* of the system (see also Ely & Plomp [86]).

In this context it is important to draw attention to the term *educational technology.* Often its meaning is only associated with the delivery media. However, today educational technology encompasses a much broader definition: it is a method for solving educational problems. A systematic or systems approach consists of the following general phases: 1) Definition of the problem; 2) Analysis of the problem and selection of a solution; 3) Development of the solution; 4) Implementation and testing of the solution; 5) Evaluation and revision (Romiszowski [81], Ely & Plomp [86]). Such a technological, problem-oriented approach is the necessary and only sound basis for successful programs and systems.
2.4 Towards an application: some general considerations and guidelines

This section discusses key issues related to the application of tele-education systems. First, it is important to distinguish various tele-education network configurations. Second, a summary of the analysis and design process steps is given, based on the educational technology approach. Finally, a short introduction is given to the technological building blocks of such a system, as far as it concerns computer related issues.

2.4.1 Classification of configurations

Beside the conventional applications of tele-education systems, various authors suggest new possibilities where tele-education could offer a solution (Dunnett [85], Cerich & Jallade [86]). Dunnett [90] distinguishes three quite separately configured ways of tele-education called: hub-school, hub-teacher and networking. However, a fourth configuration should be identified, which we call extension(s). See Fig. 2.3.

a) Hub-school (area-school). This is where a school (or university) is given extra resources and teachers so as to be able to conduct programs for the surrounding small schools. It will also include the preliminary years for those students who may proceed on the hub-school for a final year of specialized study. A group of schools can similarly share expertise by each in turn taking the lead for different subjects. Such a hub- or area-school may also assist teacher development conferences in order to supplement the limited resources available in a single remote school.

b) Networking. This is a situation in which groups of small schools pool their resources and share the provision of an experienced or specialized teacher as required.

c) Home-based (hub-teacher). This is a supplement to, an extension of, or sometimes a complete replacement of, correspondence where a teacher can be provided to satisfy a distant need by communication technology.

d) Extension(s). In this case the school establishes an extension at one or more remote locations, where the course program can be followed by tele-education. Hybrid solutions are possible: tele-education supplemented by face-to-face lecturing.
Fig. 2.3 Configurations for tele-education ($S =$ Student(s); $T =$ Teacher)

The optimal configuration (this could be a combination of the basic configurations mentioned) depends on the local situation, the target groups, and the objectives and criteria of the educational institute.

2.4.2 Educational objectives and criteria

In accordance with the argumentation of paragraph 2.3, it is necessary to start with a careful analysis of the educational problem wherever a tele-education solution is considered. Before looking at possible technologies, the objectives should be clearly formulated, the target groups should be identified and the user requirements should be generated. Besides, it is necessary to analyze the context in which the system should function concerning educational, organizational, technical and budgetary aspects. Although an iterative approach is not excluded, the basic principle should be that this problem analysis and definition process is being finalized before entering the design phase. Following the general instructional design approach of Romiszowski [81], the next discussion may be considered as a summary set of guidelines through this process of problem definition and analysis.
Definition of the problem
The educational needs of (potential) learners might be met inadequately or not at all. Define these needs and problems. E.g., 'A certain geographical isolated community has no access to the continuing education facilities of a university due to the lack of a distance-teaching service'; 'A certain rural school is not able to teach a number of specialist courses due to shortage of certified teachers.'
State what exists and what should be achieved in precise and measurable terms, but avoid statements which imply specific solutions. State the 'real-life' objectives.
Consider the probability of success in developing a solution. Are there any environmental constraints which make it improbable/impossible to achieve some objectives? Estimate the probable cost of a solution. Estimate the relative worth of solving this problem, as opposed to using the same resources for some other problems (priorities).

Analysis of the problem and selection of a solution
Identify the target group(s) and analyze their educational needs. Questions which should be answered: Are the members of the target group living scattered over a wide area or are they concentrated at a few locations? What are their constraints concerning available time? What are the entry levels? Are they already participating in existing course programs or not? What is the expected level of motivation? Are the level and type of the needed education specified?
Design, in general terms, all solutions which seem possibly appropriate. Identify the resources required for developing, producing and implementing each solution. Consider the advantages and disadvantages of alternative solutions in terms of: effectiveness, cost, probable development difficulties, probable implementation difficulties, possible control difficulties, other factors considered important (e.g., political, cultural), possible side effects.
Broad design issues are: is the solution an extension of existing methods or does it require adoption of complete new teaching methods, media, and/or organizational consequences?
Specific issues concerning distance teaching solutions: interactivity, real-time capabilities, hybrid alternatives (courses partly realized by face-to-face teaching), learning effectiveness.
Specific issues concerning tele-education: configuration, telecommunication tariffs and infrastructure, equipment serviceability, computer infrastructure, user-friendliness of the system(s), adaptability to different types of users (experienced, novice, incidental), training requirements, evolvability, connectivity to other tele-education networks, hybridization (use of mail-based facilities).
An important factor to be considered in the event of adopting new technologies is the organizational environment necessary for successful innovation. There are basically two ways in which a new technology may be used within a system. The first is for it to be added to an existing system. This usually means that it does not play a central role and adds only cost to the system. The second is for it to replace an existing technology. This normally means not only changing technologies but also methods of working. Then, changes in technologies will need to be accompanied by structural changes within the organization, as well as changes within individual work practices, to ensure that resources and decision-making powers are up to the requirements of the new technology. This is usually the main barrier to innovation, since those with decision-making powers are often those controlling the resources associated with the 'old' technologies that are under threat (McDonnell [92], Bates [88], compare also Arnbak [88] and Carey [89]).

**Development**

The development of a solution involves basic course plan(s) selection, detailing of the objectives (sub-objectives), deciding on student groupings and learning strategies. During this development phase the media should be selected. Dunnett [90] distinguishes supportive and essential media: "If media are used, they must benefit something and the bonus must be identified. When media are used as a bonus they risk being sidelined and being 'in' or 'out' as various financial climates operate. In a continuous mode of distance-related education, the process simply ceases without medium, as the media are part of the process; consequently they immediately become an essential element. This clearly is the difference of the 'distance' situation which creates the need for media as being essential".

Fig. 2.4 outlines the decision process with respect to the application of media and tele-education, in general terms. It emphasizes that it should be based on educational objectives and systematic analysis of alternative solutions (taking into account resources and constraints), rather than 'technology-driven'. However, for those situations where a tele-education solution is considered, it is worthwhile to draw attention to the available delivery media. Besides conventional media like television, telephone and radio, new media evolve based on more advanced telecommunications and computer technology. Therefore, assuming the reader to be familiar with the former, in the next paragraph a summary is given on telematics and multimedia technology.
2.4.3 Telematics and multimedia

2.4.3.1 Introduction

The word 'telematics' signifies the merger of information services, information technology and telecommunications. 'Information' includes human language and logic, and is normally the content in a telematic message. 'Information technology' includes computers, their central and peripheral components, and terminals (Ryan [81]).

One of the most prominent exponents of this technology is electronic mail (E-Mail). Today, via computer networks, electronic messages can be exchanged between persons separated from each other in space and time without the delays of postal delivery. Still nascent, this mail net already encircles most of the globe (Perry & Adam [92]). Related services are conference groups and bulletin boards, enabling users to communicate at group level. Further, the possibility to access remote computers offers facilities like database search and retrieval, software and data exchange, and tele-computing. All these services have been recognized early on as a benefit for distance education (Ryan [81],
Cross [83]). However, at the end of the Eighties, multimedia technology came into view, giving new and promising perspectives to these predominantly text-based services.

Multimedia systems involve the generation, processing, storage, transmission and presentation of information consisting of text in combination with graphics, audio, animation and video. Another, less technology-oriented definition is given by Jagerman [91]: "Multimedia can be defined as the technology that enable applications to present data to or get data from users, directly or via other applications, by modulating the form of the presentation of this data to one or more sensory channels of the user, thus permanently responding to the expressed perceptual needs of the user". The following media are distinguished from a technical point of view:

a) **Video.** This data represents full-motion video, in analog or digitized form. The advantages of digitized video are that such data can be transmitted via digital networks, and can be stored as part of a digital database.

b) **Low frame-rate video.** Since full-motion video demands significant bandwidth or storage requirements, for some applications it may be necessary to apply a low (∼ 1 Hz) or ultra-low (< 1 Hz) frame-rate. The latter are known as slow-scan or freeze-frame TV methods.

c) **Animations.** Animations can be seen as a sequence of computer-generated pictures, which can be displayed to give the illusion of full motion. These data may be object-oriented and contain various hierarchical levels.

d) **Geometric pictures.** Pictures composed of standard graphical objects, usually organized in layers and groups. Frequently these representations can be zoomed in or out.

 e) **Raster-scanned pictures.** Digitized photographs, where each picture element (pixel) may be represented by two or more colours. An important fidelity criterion is the resolution (the density of picture elements).

f) **Handwritten graphics.** Data obtained from an electronic writing tablet with a handheld stylus. The data represents the dynamical movements of the pen on the tablet with selectable attributes like colour, thickness and style.

g) **Sound.** Recorded audio data, either in digitized or analog format, representing voice, music, synthesized, signalling or environmental sound.

h) **Text.** Unstructured or formatted text.
2.4.3.2 Application categories

Multimedia applications can be distinguished in three main categories:

1) Information systems
   This encompasses the preparation, storage, cataloguing and retrieval of multimedia
   information and documents. Electronic documentation offers easy and fast informa-
   tion access, saves weight, space and environmental resources, and enables fast
   distribution. In combination with multimedia, better comprehension with reading
   and training can be achieved. Besides conventional hierarchical, sequential and
   word-based index methods, multimedia information systems often feature search
   methods based on hyperlinks, i.e. navigation through information via predefined
   relations between objects in different (parts of) documents ('hypertext'). Library
   services, point of information systems, maintenance data, training courses, referen-
   ces, presentation and visualization systems, are just a few examples of the many
   possibilities.

2) Communications
   This includes store-and-forward message handling: multimedia E-Mail and file
   transfer (Borenstein [91]), remote access to multimedia databases, and real-time
   conferencing. The latter is basically audio or video in combination with a common
   text and graphics work area to improve communication and cooperation between
   persons separated by distance. Real-time conferencing may be applied for telecon-
   ferencing including multipoint meetings, tele-education and training, and computer
   supported cooperative work (CSCW).

3) Human Computer Interfaces (HCI)
   Voice commands, spoken help, and animations support are among the multimedia
   applications which may improve desktop productivity of word processors, spread-
   sheets, data bases and calendars. Multimedia can play a role in the application of
   direct manipulation interfaces (DMI). In a DMI, the user interaction is performed
   by manipulating objects (e.g., icons) belonging to a consistent system on the screen,
   instead of giving text commands (Jagerman [91]).

2.4.3.3 Standards and communications

Various standards involving multimedia issues already exist or are being developed now.
To support efficient storage and transmission of multimedia information, image and data
compression methods are used. Three digital video standards that have been proposed are the Joint Photographic Experts Group (JPEG) standard for still-picture compression; the Consultative Committee on International Telephony and Telegraphy (CCITT) Recommendation H.261 for video teleconferencing; and the Moving Pictures Expert Group (MPEG) for full-motion compression on digital storage media (Ang, Ruetz & Auld [91]). Audio may be represented at various quality levels, e.g. the CCITT standard for 7 kHz audio (G.722, Jayant [90]). The Office Document Architecture (ODA) and the Standard Generalised Markup Language (SGML) are two international standards developed for the representation and interchange of documents between different platforms (Brown [89]). ISO’s Multimedia and Hypermedia Information Coding Expert Group (MHEG) is establishing standards for real-time interactive communicating applications. These standards aim at defining a common base for many of the multimedia and hypermedia applications which will be developed in the coming years in different fields. This common base is a coded representation of independent and elementary units of information, which will be specified as objects, and handled or interchanged by the different applications making use of them (Kretz & Colaitis [92]).

Besides the necessity of interchange formats, the transmission of multimedia information involves standardization areas concerning bitrate capacity, network delays and synchronization issues. The synchronization can be provided by the network, the system services or the application itself. While Local Area Networks (LAN) like Ethernet only offer asynchronous communication, more recent systems like Fiber Distributed Data Interchange (FDDI) also offer the possibility of isochronous and synchronous links. Wide Area Networks (WAN) based on packet switching generally cannot meet these requirements completely. However, Broadband-ISDN, based on the ATM (Asynchronous Transfer Mode) concept, has the potential to support the transfer of multimedia satisfactorily and flexibly.

2.4.3.4 Platforms and systems

The advent of optical storage systems was a catalyst for the development of multimedia technologies. A high storage capacity and data rate, in combination with random access and data integrity, were the key factors resulting in digital audio, digital video and various other bulk data storage applications.

Digital Video Interactive (DVI) technology was introduced for compressing and presentation of real-time video sequences on CD-ROM. The storage capacity of one CD-ROM disk corresponds to 72 minutes full-motion video. A typical DVI end-user system consists of a PC containing a DVI add-on board and accompanying driver software. The
peripherals should include a CD-ROM drive, and an audio system (Fox [89]).

CD-Interactive (CD-I) is also based on the storage of all information on one CD. But while DVI is pushed as the digital video technology for the professional market, from the beginning on CD-Interactive was intended for consumer applications. CD-I includes CD-audio and full-screen full-motion video based on the MPEG standard. For consumer use, a 'Base Case' is defined, representing a player with video output and a trackball as the only input medium for user interaction without the need for additional equipment like a PC with a CD-drive. However, emerging standardization activities indicate that the DVI and CD-I worlds are gradually converging to each other: CD-I will become a more open technology and on the other hand dedicated DVI players enter the market. The strength of both technologies from the user's point of view is that highly interactive multimedia applications can be realized (Frank [91]).

Besides these retrieval-based methods, multimedia message and real-time communications become an increasingly important issue. Recently a number of research projects have been carried out involving the development of multimedia teleconferencing systems, mostly aimed at ISDN, B-ISDN or LAN type networks, and based at workstations or high-end PCs (Robinson, et al. [91], Clark [92]). Although in most cases tele-education is mentioned as one of the possible applications, these systems are developed primarily for (multipoint) conferencing and group work (CSCW). Nevertheless it is probable that in the near future multimedia communication systems will play an essential role in the field of tele-education technology (e.g., DeSio [92]).

2.5 Summary and conclusions

This chapter introduced distance education. Reasons for the implementation of distance education facilities may be geographical barriers, time constraints or other personal limitations which prevent following regular face-to-face courses. Distance education suffers a number of drawbacks of which the long response time is the most serious one. However, during the last decades the advent of modern telecommunication technologies has changed the prospects of distance education auspiciously. Telecommunication enables real-time links and interaction, eliminating a number of the typical disadvantages of distance education. Advancing technologies drawn from informatics, micro-electronics and satellite communications have stimulated a rapid development of these new educational media in recent years. When telecommunication plays an essential role in a
distance education system we use the term *tele-education*. By definition, tele-education encompasses those forms of distance education employing telecommunication technologies.

We have given a classification of tele-education methods, involving two main categories: mail-based and real-time tele-education. Mail-based methods employ computer mediated teaching based on electronic mail services. In comparison with the conventional correspondence course, the response times are reduced significantly and mutual student interaction is made possible. However immediate feedback is only possible by real-time methods employing audio or video variants in most cases. Furthermore an overview was presented of the various tele-education methods implemented in real educational situations, including a number of prominent examples. Described were one-way systems, two-way audio, two-way video, two-way graphics & text with two-way audio, slow scan video with two-way audio, electronic class, and virtual class.

When considering the application of new technologies in education, the importance of a systematic and problem-oriented approach has been stressed. The misconception that technological innovation in itself guarantees improvement, has negatively influenced a large number of promising projects, with sometimes even a complete failure as a result. In this context the significance of educational technology has been emphasized. This involves a systems approach consisting of the following general phases: 1) definition of the problem, 2) analysis of the problem and selection of a solution, 3) development of the solution, 4) implementation and testing of the solution, 5) evaluation and revision.

Finally, some general considerations and guidelines were presented which are of importance to the application of a tele-education solution: various distinguishable tele-education network configurations, a summary of the analysis and design process steps based on the educational technology approach, and a short introduction to telematics and multimedia, resembling the computer-related building blocks of such a system.
References in Chapter 2

Algra, T. [88], "Towards a tele-education network on the Netherlands Antilles and Aruba", University of the Netherlands Antilles, UNA-Cahier, No.27, (in Dutch), Curacao, 1988


Arnabak, J.C.[88], "Trans-border communications networks: technology-driven or social innovation?", European Management Journal, Vol.6, No.4, pp.330-337, 1988


Baldwin, L.V. [84b], "Instructional television", IEEE Spectrum, pp.110-122, November 1984


Bates, A.W. [88], "Technology for distance education: a 10-year prospective", Open Learning, pp.3-12, November 1988


Borenstein, N.S. [91], "Multimedia Electronic Mail: will the dream become a reality?", Communications of the ACM, Vol.34, No.4, pp.117-119, 1991

Bouck, G. [87], "Distance education using satellites", Proc. 3d annual gen. conf. CANTO, 1987


Centini, B.A. [86], "Educational opportunities through telecommunication", Second Annual Conference of the Caribbean Telecommunications Council, San Juan, 1986


Champness, B. [90], "Eurostep and Olympus, Progress so far", Proc. of the 1990 Eurostep Conf., Lisse, Netherlands, pp.11-14, 4-6 April 1990

Chaplin, J.[89], "ESA’s Olympus satellite and distance learning in Europe - The creation of a European User’s association (Eurostep)", ESA Bulletin, No.60, pp.63-67, 1989


Cross, T.B. [83], "Computer teleconferencing and education", Educational Technology, pp.29-31, April 1983


Dunnett, C.W. [90], "Open access - open media", Educational Media International, Vol.27, No.4, pp.190-207, 1990


Guiton, P. [84], "Use of an FM radio subcarrier for university distance education", *Progr. Learning and Educ. Technol.,* Vol.21, No.4, pp.310-313, 1984


Jayant, N.S. [90], "High-quality coding of telephone speech and wide band audio", *IEEE Communications Magazine, Vol.28, No.1*, pp.10-20, 1990

Kretz, F., F.Colaitus [92], "Standardizing hypermedia information objects", *IEEE Communications Mag., Vol.30, No.5*, pp.60-70, 1992

Lalor, G.C. [84], "The University of the West Indies Distance teaching project (UWIDITE)", *Higher Education (L.A. and the Caribb.),* No.15, 1984

Lalor, G.C., C.Marrett [86], *Report on the University of the West Indies Distance Teaching experiment*, Kingston: University of the West Indies, 1986


Perry, T.S., J.A.Adam [92], "E-Mail pervasive and persuasive", *IEEE Spectrum, Vol.29, No.10,*


Rosenthal, L. [87], *The Eastern Caribbean Center Telecommunications Project*, University of the Virgin Islands, February 1987


3 REQUIREMENTS FOR A REAL-TIME MULTIMEDIA NARROWBAND TELE-EDUCATION SYSTEM

3.1 Introduction

Geographical barriers often prevent certain communities to benefit from the educational facilities of regular universities and schools. This chapter deals with specific situations where real-time tele-education is an effective solution to this problem. Looking at a case representing a broader range of similar situations, the problem areas are analyzed and various approaches to a solution are studied. The adopted systematic problem-solving methodology is inspired by the approach advocated in Romiszowski [81], namely, the problem analysis leading to the application of a particular technology instead of the other way around (compare also Ely & Plomp [86]). In section 3.2 a requirements analysis is performed of a multimedia real-time tele-education system, while section 3.3 proposes a conceptual system architecture and identifies a number of critical technologies.

3.1.1 Problem definition

In a substantial number of regions the facilities of the existing educational institutes cannot be used by all qualified segments of the population. Reasons may be the geographical distribution of the population or topographical barriers, preventing people from following regular face-to-face courses. The Caribbean Basin, with a multitude of small island communities mostly without any form of higher education, is a typical example of such an area. Consequently, today universities such as the UWI (University of the West-Indies with main-campuses on Jamaica and Trinidad) and the UVI (University of
the Virgin Islands) operate successful tele-education networks (Lalor & Marrett [86], Rosenthal [87]). Another example, which we shall use as a case in this chapter, is the University of the Netherlands Antilles (UNA) located at Curacao. The UNA is the only institute of higher education on the archipelago of the Netherlands Antilles, consisting of five small islands with mutual distances up to 800 km.

**Table 3.1 Netherlands Antilles And Aruba**

<table>
<thead>
<tr>
<th>Island</th>
<th>Population</th>
<th>Area (km²)</th>
<th>Capital</th>
<th>Higher Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aruba</td>
<td>60,312</td>
<td>190</td>
<td>Oranjestad</td>
<td>UA</td>
</tr>
<tr>
<td>Bonaire</td>
<td>9,753</td>
<td>288</td>
<td>Kralendijk</td>
<td>UNA</td>
</tr>
<tr>
<td>Curacao</td>
<td>147,388</td>
<td>444</td>
<td>Willemstad</td>
<td></td>
</tr>
<tr>
<td>Saba</td>
<td>965</td>
<td>13</td>
<td>The Bottom</td>
<td></td>
</tr>
<tr>
<td>Statia</td>
<td>1358</td>
<td>21</td>
<td>Oranjestad</td>
<td></td>
</tr>
<tr>
<td>St.Martin</td>
<td>13,156</td>
<td>34</td>
<td>Philipsburg</td>
<td></td>
</tr>
</tbody>
</table>

*Fig. 3.1 The islands of the Netherlands Antilles and Aruba*
### Table 3.2 Enrolled students UNA 1987-1988 ('auditors' between brackets)

<table>
<thead>
<tr>
<th></th>
<th>Law</th>
<th>Social &amp; Economic Sciences</th>
<th>Technical Sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aruba</td>
<td>3</td>
<td>25</td>
<td>17</td>
</tr>
<tr>
<td>Bonaire</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Saba</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Statia</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>St.Martin</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

The countries of the Netherlands Antilles and Aruba belong to the Kingdom of the Netherlands. The islands of Aruba, Bonaire and Curaçao near the coast of Venezuela accommodate the major part of the population, with Papiamento as their native language. The people of the smaller islands of St.Martin, Saba and Statia, some 800 kilometres to the north, are speaking English. However on all the islands Dutch is still the official language in education, government and law. From Table 3.1 it is noted that the islands are characterized by small communities with a medium to high population density (Palm [85]). The various islands possess a relatively well-developed infrastructure concerning telecommunications, roads and utilities. Primary and secondary education is organized according to the model in use in the Netherlands.

The UNA is a small university consisting of three faculties: Law, Social & Economic Sciences, and Technical Sciences. The latter faculty is divided in the departments Electrical Engineering, Mechanical Engineering, and Construction & Civil Engineering. Table 3.2 lists the numbers of students enrolled during the academic year 1987-1988, for each of the faculties, divided according to native islands (Algra [88]). Values between brackets represent students following a limited course program, in order to upgrade their knowledge in a specific field without the desire to obtain a grade: they are known as auditors or hospitants. Students from other islands than Curaçao usually reside at the University Campus. A part of the students from Curaçao (estimated 18%) belongs to the category of employed adults studying part-time at their own pace. Obviously, inhabitants of the other islands who are bound to their place of residence due to a job or other personal obligations, are excluded from these facilities. The same applies for potential auditors. A significant observation is that only 13% of the students are native of Aruba, Bonaire or the Windward Islands, while these areas represent 37% of the total population. Apparently, probably over 150 students who should be serious candidates for a study on the UNA (except Aruban law students), decide to apply at a university abroad or do not leave their island at all. Studying in Europe or North America may be attractive, but this custom contributes to a steady brain drain since a substantial part of these students does not return to their native country. These figures indicate that the other islands benefit less than Curaçao from the national university. In addition to its graduate
and under-graduate programs the UNA offers various extra-mural and post-graduate courses. Moreover, at the university numerous activities are organized like seminars, symposia, and lectures for the general public. Obviously, only the population of Curaçao profits by these services. This is emphasized by the following quotation of a Netherlands Antillean governmental report stating (Comm. [83]): "The complaints concerning the limited area covered by the UNA's educational facilities are well-known, and given the geographical reality of the Antilles, it is not surprising that a solution is searched in the direction of distance education. However extra-mural education, whether carried by distance media or not, is not yet an Open University, however to our opinion it should constitute an inherent part of the course facilities of the UNA".

Before Aruba left the Netherlands Antilles to become a separate country within the Kingdom of the Netherlands in 1986, the UNA operated an extension at this island for Aruban law-students. Today Aruba has its own university: the UA (University of Aruba). For some time UNA's law faculty has organized a limited number of undergraduate courses at St. Martin consisting of face-to-face sessions in combination with audio-tapes of lectures recorded at the Curaçao campus. However, high costs, together with low numbers of participating students, did not justify continuation.

The case of the UNA is representative of a broad range of situations where a significant part of the population does not benefit by the educational facilities of the universities present in the country due to geography. As a result some areas are characterized by lower average levels of education and, in relation to that, a lagging technological and economical development. It should be kept in mind that such educational facilities are not only directed to students following full-time face-to-face courses but may be aimed also at other educational segments like continuing education, professional training, and lectures or symposia for the general public. Two other main characteristics of cases similar to the UNA are the limitations typical of the circumstances in a developing country and the small scale of the institute. The former may include several constraints and limiting conditions varying from a low level of equipment serviceability to an insufficient educational infrastructure. The latter has its impact on the available resources.

The educational problem outlined for the Antilles requires a solution involving the facilities of the UNA to become accessible for inhabitants of all the islands, resulting in:
1) A larger number of students of the islands (other than Curaçao) enrolling in the university's graduate and undergraduate programs
2) A regular and significant participation of inhabitants of these islands in seminars, symposia, lectures, etc.

Generalizing, the main objective of a solution to the presented problem is: a significantly higher level of participation of inhabitants from the 'problem' areas in the university programs and activities offered.
3.1.2 Analysis and selection of a solution

In this paragraph we will consider the various target groups and identify possible solutions that may be applicable. These solutions will be compared according to various proper selection criteria. We will not refer to the Antilles-case explicitly, but analyze and describe the problem in more general terms, although we are still dealing with a class of educational situations of which the Antilles is a representative example. For convenience isolated remote communities will be referred to as 'islands'.

Two primary target groups can be identified:

1) University students
   These are students who fulfil the admission requirements of a course program at the university to obtain a grade. Their level of motivation is not expected to be different from their peers at the university campus. Part-time students and auditors also belong to this target group.

2) Selective professional groups
   Post-graduate and continuing education in the form of symposia, seminars and courses are mainly directed to professionals in order to upgrade their knowledge and skills. Some courses may be organized for the personnel of a particular company, institute or government service. Hence the size and the composition of these course groups is quite variable, but in general the level of motivation and concentration will be rather high.

Although the spin-off of a particular solution could probably include facilities for the general public, this should not be considered the primary target group, similarly to the regular (face-to-face) situation. Furthermore, a typical characteristic for the various target groups on all the islands is that a central educational facility on any of the islands can be situated within convenient travelling distance, due to the reasonable road conditions and the small areas of the islands.

With respect to the problem defined, four main approaches to a solution may be adopted: 'fly-in' courses (face-to-face), conventional distance education, E-Mail tele-education and real-time tele-education. Although further detailing of a solution may reveal the attractiveness of hybrid variants, initially we will compare just these four alternatives on a selection of proper criteria.

a) Fly-in courses (face-to-face teaching)
   Description: This solution offers (a part of) the course program of the university realized at the various islands in its unaltered original form. Frequent travelling of
university teachers between the islands is unavoidable, although incidentally local part-time teachers may be contracted. At all islands served, one properly equipped building should be available with sufficient class-rooms and facilities. This solution has been applied in very few cases, and generally cannot be justified due to the low numbers of participating students.

*Advantages:* The solution offers all the known benefits of face-to-face teaching such as a maximal level of interactivity, immediate feedback, and better student-coaching possibilities. No adaptation of curricula and teaching methods is required, which is profitable for both students and teachers. In general the method allows a fast and flexible implementation.

*Disadvantages:* Travelling expenses are high. Moreover, frequent travelling will reduce the effectiveness of available human resources significantly. Further the method is not feasible for guest lecturers from abroad with their generally limited time budgets. Obviously simultaneous teaching of multiple classes or groups by one teacher is not possible.

b) *Conventional correspondence courses*

*Description:* This is the traditional correspondence course, extended with audio or video-tapes where appropriate. It requires a complete redesign of the courses offered. Institutes such as the UNA do not possess the experience and knowledge necessary to implement this type of education. Hence qualified courseware designers and teachers should be contracted to perform the job.

*Advantages:* The correspondence course approach is a 'proven technology' involving rather accurate projecting of resources and expenses, and reliable prediction of course performance. Another advantage is that a central facility on the islands involved is not required.

*Disadvantages:* The solution requires a completely other type of educational organization, and extra resources for courseware development. It suffers from the well-known disadvantages of conventional distance-education such as a long turn-around time due to postal delivery of materials, a low level of interactivity and the problem of pacing and motivating the participating 'remote' students. Moreover, the correspondence course is not an effective vehicle for efficient time-compressed professional courses and seminars, and not suitable to replace guest-lecturing.

c) *E-Mail tele-education*

*Description:* This solution involves correspondence courses applying E-Mail delivery of educational materials, assignments, questions and answers, possibly
extended with services like conferencing, bulletin boards, and telecomputing. The turn-around time as compared to postal delivery is significantly reduced. It should be noted that this method is based on individual learners rather than groups, thus requiring a quite large base of terminal facilities. At the university ample computing capacity should be available for operation of such a tele-education network and services.

Advantages: The solution uses proven technology; E-Mail networking and services are broadly used, also for tele-education purposes. Mutual student interaction is possible by conferencing and bulletin boards, and the method allows telecomputing, software downloading and remote database retrieval, e.g., for library services.

Disadvantages: A certain amount of training is required to be able to use the services. Except for the turn-around time, all other disadvantages of the correspondence course still apply (resources, cost). Graphical correspondence and in particular multimedia-type services are still in a development phase (Arnbak et al. [89], Borenstein [92]). Central facilities should be situated at the participating islands with sufficient terminals (PCs and modems) for students who cannot afford private equipment. At these reception sites a facility manager should be responsible for operations and maintenance.

d) Real-time tele-education

Description: This involves remote teaching of classes or groups of learners in real-time, with sufficient feedback possibilities (interactivity) and adequate media. In general the course plans do not need updating, although minor modifications of course contents might be recommendable due to media limitations. A two-way audio system with a one-way visual component is a minimal requirement for implementation. At the participating islands a reception site needs to be established including adequate classrooms.

Advantages: This type of distance education is the best replicant of normal face-to-face teaching, meaning that teaching methods and curricula require only minor modifications, and immediate teacher-student and student-student interaction remains possible. Multipoint networking and candid classroom methods allow simultaneous teaching of multiple learning groups at the university campus and the involved islands.

Disadvantages: For small institutes ITV (Instructional Television) might not be affordable due to the high costs of using broadband telecommunication channels. In that case narrowband systems could be applied, of which computer-based
Table 3.3 Comparison of four approaches to a solution

<table>
<thead>
<tr>
<th></th>
<th>Learning effectiveness</th>
<th>Development difficulties</th>
<th>Implementation difficulties</th>
<th>Other factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fly-in courses</td>
<td>High</td>
<td>No (replicants of existing campus-based courses)</td>
<td>High cost due to travel and impact on human resources Qualified teachers may be scarce No multipoint and candid classroom</td>
<td></td>
</tr>
<tr>
<td>Correspondence courses</td>
<td>Moderate</td>
<td>New teaching methods to be adopted New course-plans and courseware to be developed</td>
<td>High cost due to impact on human resources Qualified teachers may be scarce</td>
<td>Not applicable for guest-lectures, seminars and symposia</td>
</tr>
<tr>
<td>E-Mail based tele-education</td>
<td>Moderate</td>
<td>Proven technology except pictures &amp; graphics New teaching methods to be adopted New course-plans and courseware to be developed</td>
<td>High cost due to impact on human resources Qualified teachers may be scarce</td>
<td>Not applicable for guest-lectures, seminars and symposia</td>
</tr>
<tr>
<td>Real-time tele-education</td>
<td>Moderate to high</td>
<td>May require development of technology (multimedia-based terminals)</td>
<td>Availability, reliability, and tariffs of telecommunication infrastructure</td>
<td>Evolvability, connectivity depends on technology chosen</td>
</tr>
</tbody>
</table>
multimedia systems offer the best prospects to realize effective real-time tele-education. However multimedia tele-education is still a nascent technology. At the reception sites, a facility manager should be responsible for operations and maintenance.

Table 3.3 presents a schematic overview of the above considerations. It can be concluded that face-to-face courses, correspondence courses and E-Mail based tele-education are not appropriate solutions for the envisaged situation:

a) Correspondence courses and E-mail based teaching methods do not guarantee sufficient learning effectiveness with respect to the envisaged target groups. For these forms of education a level of motivation and concentration is required which cannot be expected from undergraduate students.

b) Courses for professionals (often presented by guest lecturers used to traditional face-to-face education) should allow time-efficient teaching which is only possible by real-time interactivity.

c) Non real-time distance education requires a completely different type of organization, and other teaching methods. New methods would require extra human resources both for the development and the operational phases, which is not feasible and affordable for a small-scale university.

d) In this context it should be noted that a frequent problem of developing countries is scarcity of qualified experts (i.e., certified teachers and courseware developers).

e) It should be taken into account that the introduction of a novel technology simultaneously with new working methods implies a higher threshold to get the system introduced and accepted (Bates [88], Arnbak [88], Carey [89], Arnbak et al. [90]).

The remaining part of this study will be devoted to the design of a technical infrastructure necessary to realize a real-time tele-education service in the envisaged situations. In this context, the following considerations should be taken into account:

1) A workable operational solution does not only consist of a technical infrastructure, but it constitutes also well-designed educational and organizational methods, procedures, and facilities. As a follow-up, a pilot experiment is recommended covering all these aspects (refer to 9.2).

2) From our 'global' analysis it was concluded that real-time tele-education should serve as the 'main vehicle' for the distance teaching process. But other, supplementary methods are not excluded. A more detailed analysis may reveal the necessity of hybrid variants. This might involve real-time tele-education courses
completed with a limited number of face-to-face sessions, additional E-Mail services and postal delivery of course materials. Thus, in addition to the regular tele-education program, teachers would visit the islands periodically for the purpose of face-to-face lectures, personal coaching and remedial teaching. Furthermore, hybrid solutions might be applied in another sense, involving particular phases of the (under)graduate program to be attended on campus. This could be necessary for laboratory instructions requiring the use of expensive specific equipment or tools. The first year of certain undergraduate courses might be taken completely ‘on distance’, while after successful completion the remaining phases with more specialization would require on-campus studying (Lalor & Marrett [86]). Conversely, during the final part of the study the thesis research project on its turn could be performed at the native island with student-coaching via the tele-education network. This allows a direct assessment of specific problems of remote communities and mobilizes university knowledge for the benefit of these areas.

The real-time tele-education solution should be implemented primarily as an extension configuration. Still, cooperation with other universities by networking is an attractive additional possibility. For this purpose sufficient flexibility and extensibility of the system should be guaranteed.

Realization of networking involving the UNA and the UA is extremely attractive due to the great similarity between the curricula and the organization of the two law faculties. The university of Surinam and certain institutes in the Netherlands are also potential candidates for such a network.

In the next paragraphs the requirements for a real-time tele-education system are analyzed and defined, and a possible system concept is developed.

3.2 Requirements and constraints

3.2.1 Limiting conditions and constraints

The application of a real-time tele-education system implies a number of specific conditions and limitations concerning educational, financial, technical and organizational aspects as relevant in the envisaged situation. Again these aspects will be considered with respect to the Antilles case, representing a large group of similar situations.
3.2 Requirements and constraints

a) Educational aspects

An important constraint is that the system should be appropriate to extend existing teaching methods. New methods are not feasible because of investment and operational cost, scarcity of qualified experts, and the (probably) high acceptance threshold. Thus, a technology should be selected enabling continuation of the current teaching methods as applicable in face-to-face teaching (McDonnell [92]). Only minor adaptations of these methods or the curricula are acceptable. Once a new technology is successfully applied and accepted, after a certain time more efficient and advanced teaching concepts might be considered. For instance, a gradually increasing use of conventional and/or innovative self-presenting media (e.g., printed materials, audio- and video-tapes, interactive video, simulation programs, CAI programs). However in the context of this investigation we will only study methods allowing teaching methods close to face-to-face education.

b) Technical aspects

In most countries the telecommunication infrastructure is in a process of rapid development, resulting in a steady progress of services, quality and capacity. However, the envisaged tele-education system should also function properly in situations of minimal telecommunication possibilities. In practice this means that in many relevant cases only analog dial-up circuits are available, with a transmission quality allowing a maximal bit rate of 1200 bit/s. In other areas higher capacity links might be available, e.g., 9600 bit/s (analog telephone lines of international quality [CCITT, Rec. G.153], allowing advanced modems) or even ISDN (2 x 64 kbit/s, IEEE [91]). Therefore a development based on analog links of 9600 bit/s as a design driver is justifiable, but with the condition that the system should offer enough flexibility to cover a broad range of bitrates (from 1200 bit/s to 64 kbit/s).

Currently at the Netherlands Antilles and Aruba all telephone exchanges are digital including most trunk links. The telecommunication links (microwave) between the various islands also have been digitized recently, except the sea-cable between the Leeward and the Windward islands. The latter provides analog communication circuits allowing 9600 bit/s connections. The local telephone networks (i.e. the subscriber loops) for the time being remain analog and allow only reliable connections up to 1200 bit/s. In general it is possible to obtain conditioned (leased) lines for 9600 bit/s at most locations. One aspect to be kept in mind is the probability of a decreased link quality and of incidental link dropouts due to extreme weather conditions during the rain season. In other words an application should incorporate a robust communication system with built-in flexibility regarding the available bitrate capacity. Concerning the possibilities for Instructional Television (ITV), it is observed that the telecommunication infrastructure does not provide the necessary broadband channels. A dedicated microwave link would
be possible (except between the Leeward islands and the Windward islands), but would require relatively high investment and operational costs.

Another important issue is the available level of serviceability concerning the equipment used. Especially the smaller islands lack the infrastructure guaranteeing adequate service and maintenance of complex dedicated equipment. Therefore the use of standard equipment such as PCs and accompanying peripherals is recommended. To the maximum extent possible, functionality should be realized by software.

Power outages of half an hour or more are not uncommon in the envisaged areas. Technical and organizational methods should be incorporated to reduce the impact on the educational process as much as possible. For instance power backup generators and procedures for rescheduling of lecture programs could be applied.

c) Budgetary aspects

The small scale of the university ( < 1000 students) and the small target groups require a low-cost solution, in terms of both investment and operational costs of equipment and telecommunication links, and of the number of staff required (teachers, technicians, and management). With respect to equipment preferably existing microcomputers, peripherals and local networks should be used where possible.

At the UNA currently all teaching staff has a PC at their disposal. Furthermore sufficient local network facilities and PCs are available for student use. Hence the envisaged tele-education network preferably should be based on PCs. This corresponds to the condition of sufficient serviceability. Taking into account the currently installed base of PCs, the system should be developed for 286 (and higher) processors.

The development of a tele-education system appropriate for processors including the rather moderately performing 286 CPU seems debatable. However it should be born in mind that the projected installed base (world wide) of PCs in 1996 will still include 33 millions of 286 PCs corresponding to 17 percent (Byte [92]).

d) Organizational aspects

Organizational constraints are determined primarily by the small scale of the institute and the corresponding manpower budgets. Therefore implementation of a tele-education service should have a minimal impact on organizational structures, resources and procedures. A step-by-step way of implementation would be highly preferable and allows a smooth introduction of the service.
3.2 Requirements and constraints

3.2.2 Functional and performance requirements

This paragraph analyzes and defines the requirements of a PC-based tele-education system providing real-time communication. A subdivision is made according to functionality: 1) media; 2) editing & filing functions; 3) communications and configurations; 4) educational on-line features; 5) platform; 6) user interface and screen-layout; 7) performance; 8) utilities.

3.2.2.1 Media selection

Taking into account the various target groups (paragraph 3.1.2) we distinguish five types of user groups, differing more or less with respect to the use of the various media:

1) 'Class'
   This type is characterized by a small number of students (less than 20), allowing to actively participate during the lesson, by questions, answers, and discussions. The teacher should have the means to control this class interaction process. A further characteristic is that the teacher performs activities to stimulate and motivate the individual students.

2) 'Lecture'
   In general attended by a larger number of students, with a higher degree of motivation and concentration as compared to the type 'class'. The possibilities for the individual student to participate actively are reduced.

3) 'Symposium'
   The numbers of students or participants are large (e.g., more than 150). In general, the participants are highly motivated and often familiar with the subjects presented. Individual interaction is reduced to a few opportunities to raise questions or to have a panel discussion. In general, the actual lectures or presentations themselves will be not interrupted by interactive communication.

4) 'Meeting'
   These are teleconferencing applications with more than two participants.

5) 'Conference'
   Teleconferencing with two persons.

The latter two types of media-use do not concern the primary target groups. But where a tele-education system is in operation, members of the university staff and/or students
Table 3.4 Media requirements with respect to user groups

<table>
<thead>
<tr>
<th>Medium</th>
<th>class</th>
<th>lecture</th>
<th>symposium</th>
<th>meeting</th>
<th>conference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audio</td>
<td>1 *</td>
<td>1 *</td>
<td>1 *</td>
<td>1 *</td>
<td>1 *</td>
</tr>
<tr>
<td>Graphics</td>
<td>1 *</td>
<td>1</td>
<td>1</td>
<td>3 *</td>
<td>3 *</td>
</tr>
<tr>
<td>Reduced video</td>
<td>1 *</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3 *</td>
</tr>
</tbody>
</table>

1 = necessary  3 = enhancing  * = two-way mode required

2 = desirable

could use the facilities for communication, however with a lower priority. It should also be noted that the first three groups include two types of media users: the students (the actual target group) as well as the teachers.

For real-time distance education the following types of media may be distinguished: 1) video; 2) audio; 3) graphics; 4) reduced video. The term graphics will refer to all visual information not generated by a video camera. On the other hand, the term reduced video will refer to all types of video communication with a reduced frame-rate (< 25 Hz) and/or a reduced resolution as compared to standard video.

In distance teaching different combinations of these media are applied, which may be used in one-way or two-way mode. Video is an adequate medium for the presentation of educational information, for non-verbal communication and to observe the classroom situation. Normally, this medium cannot be employed due to the lack of sufficient telecommunication bandwidth. The type 'graphics' is appropriate for the presentation of educational information. Reduced video is a proper medium for non-verbal communication and to observe the auditorium. The demand for a particular medium depends on the type of media user group. Table 3.4 distinguishes the three categories necessary, desirable and enhancing and also indicates if the medium should be used in a two-way mode.

It can be concluded that the following types of communication should be offered, provided that adequate telecommunication capacity is available:

a) Audio communication

Full-duplex audio communication of sufficient quality should be possible. Provisions should be made for each individual student to be able to address the teacher.
b) **Graphics communication**

The teacher should be able to support his lectures with visual materials in a way similar to face-to-face lecturing. Also in this case two-way communication is desired. This means that the student should be able to point to or even edit the presented information on-line, implying a *shared-screen* concept.

c) **Reduced video**

Furthermore it is desirable (but not necessary) that teacher and auditorium can watch each other, with sufficient resolution and refreshment rate of pictures.

These considerations lead to the following more detailed requirements concerning the media to be applied:

**A: Audio**

Standard audio equipment should be used to mix microphone-signals from the teacher, the live class if applicable and the remote group(s). Multiple microphones in a classroom require adequate signal on-off switching, either manually or electronically. The audio-signal should be at least of (standard) telephone quality (i.e., a frequency band from 300 to 3400 Hz and S/N $\geq$ 38 dB, CCITT Rec. G.153). However, a wider signal-spectrum is preferred, e.g., 7 kHz (compare Clark [92]).

**B: Presentation screen**

This screen should be used to display the shared-screen window, system control information and the reduced-video window. The information should be clearly visible for all the students in the classroom, implying one large screen in front of the class or multiple screens at appropriate locations. The Video Graphics Adapter (VGA) display standard should be adhered to, since at low cost it provides for proper resolution (640 x 480 pixels), 16 simultaneously displayable colours out of a palette of $2^{18}$ colours and a flicker-free scan-frequency of 70 Hz. Moreover, it is a widely adopted standard for the PC, while more sophisticated video adapters usually are downward compatible to VGA. VGA enables simultaneous displaying of objects in 8 different colours (including white and black) together with black & white pictures in 10 grey levels. However the system should be compatible also with the Enhanced Graphics Adapter standard (640 x 350 pixels; 16 colours out of 64). In this (EGA) standard, only four grey levels can be used to display pictures in black & white. Hence dithering techniques should be applied to soften contouring effects. Many video adapters today support an extended VGA-mode involving the simultaneous display of 256 colours. This mode enables the display of colour pictures of tolerable quality, which is very acceptable for tele-education purposes. Preferably, the system should support this mode, too, although it should not be considered as a necessary
requirement. It should be noted that affordable VGA to NTSC/PAL (and vice versa) convertors are available today, which enable video-tape recording of lectures in a convenient and efficient way.

C: Input media

Besides a keyboard (for inputting text and control information) and a mouse (for graphics, pointing and control information) the following input media should be considered. In order to input simple sketches, diagrams, mathematical formulas, etc., an electronic writing tablet is desirable. The movements of a hand-held stylus over the tablet are immediately transferred to the PC-screen and displayed with selected attributes like colour, thickness and line-style. The alternative possibility of a light-pen is less suitable for writing on a vertical screen (McConnell [86]). Another attractive device is the electronic whiteboard, which could replace a writing tablet (Bruce & Elrod [92]).

The combination of keyboard, mouse and writing tablet form an adequate replacement of the traditional blackboard. However, to display printed or photo-copied materials (pictures, ‘overhead sheets’, figures from a book, etc.) a scanner or a document-camera should be used. The latter is a video camera vertically mounted at a document-stand, and connected to a frame-grabber card inside the PC. To display objects by reduced video like the image of the teacher, the demonstration of an experiment, etc., or even a blackboard, a second camera is recommendable.

D: Shared-screen communication

It should be possible to display text and graphics (input via mouse and keyboard), handwritten information (input via the writing tablet) and still-video pictures simultaneously on the shared screen. This window may be overlapped by the reduced-video window (see point 1).

E: Text

To ensure sufficient visibility a proper set of fonts should be selected (including shadowing attribute). The minimal character size should not be chosen too small, e.g., a size corresponding to that of videotex [CCITT Rec. T.100] is quite acceptable. This means 20 rows of maximally 40 characters. Each character may be displayed in one of the 16 colours available, at any location on the screen (horizontally oriented). Although convenient, word processing features are not absolutely necessary, considering the usually limited amounts of text on screen.

F: Graphics

Simple computer-generated graphical objects should be possible including line, circle, disc, rectangle, bar, (filled) ellipse, (filled) polygon. Besides 16 colours a number of fill patterns should be available. Other attributes are line-thickness and style (e.g., solid, dotted). These objects should be generated using a mouse-based user interface.
3.2 Requirements and constraints

G: Pictures
This involves (parts of) raster-scanned still pictures (generated by a scanner or a frame grabber in conjunction with a video camera). These items may be displayed in two colours (binary) suitable for overhead-sheet-like documents or in 10 grey levels (standard VGA). Extended VGA allows 256 colours. Editing facilities should include zoom and clip functions, while image contrast and brightness values should be adjustable.

H: Writing tablet
Handwritten information should include selectable attributes like line-style, thickness and colour. An efficient method should be adopted to handle the eye-to-hand coordination problem, which is inherent to the use of a writing tablet as a PC-peripheral device (e.g., see Mc.Connell [86]).

I: Reduced video
Preferably, the students should be able to watch the teacher. Since all the information (including the shared-screen window) has to be displayed on a single screen, the size of this video-window has to be relatively small. It should also be possible to move or suppress the window. However, in some situations the teacher aims at demonstrating something to the students requiring full-screen video (blackboard, a physical experiment, etc.). In the reversed direction, the teacher should have a real-time video image of the class at his disposal: In learning situations, visual feedback is significant, enabling the teacher to adapt to the momentary needs of the students. Although such an image can be of a reduced size and resolution most of the time, the teacher should be able to activate a higher resolution image at any moment in order to get a more detailed view of the class situation. Hence full-duplex video communication should be possible, with independently adjustable image sizes both at the teacher and the class terminals. For this purpose we specify three sizes: small (128 x 96 pixels), medium (256 x 192 pixels) and large (512 x 384 pixels), covering 20, 40 and 80 percent of the full-screen dimensions, respectively. The images should be displayed in colour, with refreshments as often as possible.

3.2.2.2 Editing and filing functions

For incidental users (e.g., a guest lecturer) it should be possible to operate the system in a rather simple way, for instance like a ‘tele-overhead projector’ using the document camera to display paper sheets instantaneously at the remote site. For more experienced users the possibilities of computer technology should be fully exploited to generate and manage effective educational presentations of a high quality level. Two ways of editing and filing of educational information should be distinguished. Off-line editing offers the teacher a way to prepare, store, re-use, and manage educational materials in advance to
the session. On the other hand, on-line editing involves the generation, modification and retrieval of objects on the shared screen during the actual communication session. Thus these modifications will become immediately effective at the remote screen. Where both types of activities require the same sort of functionality, it is advantageous to realize one application which can be used both on-line and off-line. Hence an important requirement is that editing and file management should be integrated with the actual communication functions into one application.

Often the teacher will organize his presentation in the form of a series of 'sheets' (or 'screens'). Each sheet may be regarded as an independent instance of the shared-screen and may contain objects of several types like text, graphical objects, pictures and even animations and video sequences. Such a sheet is a one-page multimedia document, which we will refer to as a page further on. The complete series of pages, forming a set of consistent educational information for the purpose of a particular lecture, will be referred to as a presentation. During the session these pages typically will be retrieved in a sequential order, although it should be possible to retrieve any of the pages at any time. Fig. 3.2 depicts the access graph of a presentation example consisting of four pages.

![Diagram showing the access graph of a four page presentation](image)

Fig. 3.2 Document access graph of a four page presentation

The requirements are listed as follows:

**A: Object editing functions**

It should be possible to move, copy, erase and modify any object of a page. Modifying may include changing colour, line-style and thickness, fill-pattern, text-font, and also zooming operations. Further the orientation on the screen should be modifiable (rotation, mirroring). For raster-scanned pictures (still-video) a clip facility should be added. Finally, the order of the objects should be alterable to enable transfer to the foreground of objects which might be overlapped by other objects input later.
B: Object-group editing functions
It should be possible to edit selected groups of objects in the same way as applicable to individual objects. Selection may be by type (text, hand-writing, graphics, pictures) and/or by a specified field (i.e., a rectangular region on the screen).

C: Miscellaneous features
Other convenient features are the adjustability of the screen background colour, a grid structure to support precise drawing, and a 'snap' facility involving the connection of certain geometric points of different objects within a selected area. For example, in this way screen-coordinates of the end-point of a line lying in the immediate vicinity of the corner of a rectangle, can be made exactly equal to the coordinates of the latter.

D: Higher level objects
Creation of frequently used higher level objects by adequate and convenient parameter-definition procedures should be possible. Examples: graph (bar or line), pie-chart, mathematical functions with coordinate axes, music scores.

E: Macro facility
The system should support a macro facility. This means that any group of objects may be defined as a macro, and added to a macro library. A macro may be called up later-on. For this purpose macros and macro libraries should be identified by user-specified names. It should also be possible to display the directory of a library, including the name and additional comments for each macro therein. For every discipline extended macro libraries may be built up to efficiently support faster generation of educational materials.

F: Presentation naming and directory functions
A presentation should be identified by a user-specified name. Any existing presentation should be selectable and retrievable by this name. It should be possible to display the directory of a particular presentation. To further support file management operations, it should be possible to display and change the current directory.

G: Page retrieval and filing functions
Page retrieval commands should include the next page of the current presentation, the previous pages and any other pages specified by the page number. A newly generated page should be added to the current presentation by either a save next command getting the highest sequence number or a save insert command with a user-specified sequence number. It should be possible to delete any specified page from the presentation sequence.

H: Page overlay
To effectively utilize the information of existing pages for the generation of a new page, an overlay command feature should be possible. This involves the addition of all the objects of a certain page (except the background) to the current page.
3.2.2.3 Communications and Configurations

The system should support multipoint configurations, in two basic modes: 1) primary/secondary mode and 2) peer-to-peer mode. The former corresponds to the intended tele-education situation where the teacher is able to admit and to inhibit alterations of the shared-screen contents by remote students. In other words, the teacher's terminal controls the editing facilities of the other terminals to a definable extent. Therefore this configuration is termed primary/secondary. A pass-word facility should protect the system from being brought into the primary/secondary mode by unauthorized persons. By default the system is in the peer-to-peer mode, in which none of the users is able to overrule the others. The latter configuration is more suitable for teleconferencing applications.

Regarding the communication processes here we will deal only with the upper three layers as defined according to the OSI reference model (MacKinnon et al. [90]). Hence the lower layers are assumed to be implemented providing reliable communication services to the upper levels. Note that these may range from simple analog telephone services and associated modems to sophisticated network services as in ISDN with advanced switching services. The choice of implementation depends on the local telecommunication-infrastructure and the associated tariffs. In all events, the application layer should be designed to include sufficient flexibility with respect to the possible bitrates. These could range from 1200 bit/s to 64 kbit/s. Furthermore it should be taken into account that in some situations a real multipoint configuration is not feasible, whereas a form of multicasting could be implemented supported by phone-in techniques for feedback. A more detailed discussion of this matter will be given in section 3.3.

The application protocol should provide the means for the user to set-up a (multipoint) connection. Call initiation, acceptance and disconnect commands should be incorporated together with visible and audible ring-up alarms. The users should be kept informed of the status of their connection(s) by a communication monitoring facility. Incidental link dropouts or late connection of one of the secondary stations should be overcome automatically, involving immediate synchronization of shared workspaces.

If reduced-video communication is activated, an automatic priority mechanism should ensure that system commands and shared-screen objects are transmitted without delay at the highest bitrate.

Multipoint teleconferencing calls for effective floor control protocols as have been observed by several authors (Aguilar et al. [86], Sakata & Ueda [90], Robinson et al. [91], Clark [92]). However in a tele-education situation such a protocol should be adapted to the specific requirements imposed by the teacher-class interaction process. The basic
principle is that only one remote group at a time is admitted to alter the shared-screen contents; nevertheless their actions may still be overruled by the teacher. Normally only that group's audio-signal is mixed with the teacher's voice signal and transmitted to all sites. Still, any group should have proper means to draw the attention of the teacher, for instance to raise a question. This should be possible by the activation of an audio-visual attention-signal, including automatic indication of the originator, to the teacher. This feature might also include a text message. Usually, the remote groups are viewing the video image of the teacher, who should be able to select between the image of one particular group, and for a 'round-robin' scheme among all the groups. If any group is given control, its video image will be displayed automatically on the screens of the teacher and all other groups. Table 3.5 lists these various video- and audio-distribution options.

Table 3.5 Options for video- and audio-distribution in a multipoint configuration with two remote groups

<table>
<thead>
<tr>
<th>Options</th>
<th>Origin of video image at terminal of:</th>
<th>Origin of audio (at all terminals)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teacher (T)</td>
<td>Group A</td>
</tr>
<tr>
<td>Teacher views A</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>Teacher views B</td>
<td>B</td>
<td>T</td>
</tr>
<tr>
<td>Teacher views both groups alternately</td>
<td>A/B</td>
<td>T</td>
</tr>
<tr>
<td>A gets floor</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>B gets floor</td>
<td>B</td>
<td>B</td>
</tr>
</tbody>
</table>

3.2.2.4 On-line realization of specific educational features

During remote lecturing various typical educational facilities are useful or even indispensable:

A: Pointing and marking

During a lecture both teacher and students should be able to introduce and move a pointer in the shared-screen window. The local and the remote pointer should be distinguishable by colour. Further a marking feature is necessary, allowing the teacher to highlight any specific part of the screen. Obviously, the pointer and the marker-generated objects do not belong to the page contents and hence must not be stored and retrieved. Both features are controlled by the mouse device.
B: Scratch facility
During lecturing a facility for an 'intermezzo-explanation' should be available, for example, when after some discussion it appears that a short improvised explanation on a certain subject is necessary. If using the blackboard and wanting the already presented information to remain unaffected, the teacher will for this purpose often use the rear side of the turnable side panels of the board. In the event that the teacher utilizes an overhead projector for his presentation, often the blackboard present in the classroom is used for the intermezzo. In the tele-education situation an equivalent facility can be realized by popping up a specific scratch window. This window can be used for inputting and editing information similar to the normal shared-screen window, with the exception of storage to a presentation file. However, the information in the scratch window should remain retrievable until the session is closed.

C: Zooming and scrolling
When using a VGA-screen as a text 'blackboard' of 20 rows of 40 characters (see 3.2.2.1), in practice situations may occur where this text field appears to be too small. Common blackboard arrangements in lecture-rooms are characterized by a height-width ratio of about 1/4, implying fields corresponding to 10 to 15 'rows' of 100 to 150 characters (taking into account the maximal viewing distance). The character capacity of the specified shared-screen window thus amounts to only some 50 percent of the standard classroom blackboard. In general this should be sufficient, since the teacher could switch between pages. Sometimes a screen of these dimensions will appear too small (McConnell [86]), for example for an extensive block diagram or a relatively lengthy derivation. The overview and the logical cohesion of the various components may be lost if the information is distributed over more pages. Therefore the following facility should be implemented. Expressed in characters, the total shared area should be extendable to a width of 80 and a height of 40. By default, the upper left quadrant is displayed. This view-port, consisting of 25 percent of the total page, can be moved to the right and/or downward. Further, by zooming out, a larger part or even the complete page may be displayed.

3.2.2.5 Platform
To reduce cost and to avoid the development and manufacturing of dedicated hardware, the envisaged tele-education system should be PC-based, as already argued in 3.2.1. The use of PCs implies several advantages:
3.2 Requirements and constraints

a) Incremental cost: The installed base of PCs at a university may be used (partly) to reduce procurement, installation and computer training costs.

b) Peripherals like frame-grabbers, scanners, writing tablets and modems are available at affordable prices.

c) A broad range of compilers, tools and libraries is available for development and support of applications.

d) The preparation of educational materials can be performed at any standard PC due to the portability of the software.

e) Educational materials can be generated, edited, re-used, stored and archived effectively as a matter of existing routine.

f) The availability of boards and parts in combination with a good service infrastructure is high. This is of particular importance in the envisaged application areas.

The system to be developed should function properly with a 286 PC provided with 1 MByte memory and harddisk (see 3.2.1). The foreseen monitor type is VGA, offering sufficient resolution and colours. However system flexibility should provide downwards compatibility with EGA type monitors. A local area network connecting the PCs of staff with the tele-education terminals would be convenient for file transfer.

3.2.2.6 User interface and screen layout

Obviously the user interface should receive considerable attention in the design of a complex PC-application. In general terms, evaluation criteria include (Nickerson & Pew [90]): what the interface, and more generally the system, permits the user to do; how satisfied the user is; and how long it takes a person to use the system at some specified efficiency level. Another criterion is how fast benchmark tests can be performed. Also important are the robustness of the system under stress, error rates and error costs. The UI types in current applications can be categorized as follows (a historically oriented overview can be found in Perry & Voelcker [89]):

a) Command-line interface

    Every command is to be input by a text string following the command prompt. Well-known examples are the command-line interfaces of the operating systems of UNIX and DOS.

b) Menu interface

    Here the user selects a command from a list of options (the menu) by one or two
keystrokes. Menu interfaces often are organized according to a well-ordered hierarchical structure. This type of interface is much more efficient (less keystrokes; reduced probability of mistakes) and easier to learn than a command-line interface.

c) **Windows interface**

For complex applications, e.g., desktop publishing, the menu interface suffers a number of drawbacks. The windows interface technique, first introduced by Apple for the MacIntosh computer, has found widely acceptance for these kinds of applications. The use of a mouse, working with icons, and the overlap of windows so that the user keeps aware of the current menu level, give this type of interface a high degree of efficiency. Its explicitly visual and intuitive nature contributes to a relatively short and steep learning curve.

d) **Pull-down menu interface**

With pull-down menus the program usually has a large workspace window with one or a few horizontal menu bars. The latter take relatively little space on the screen. Once a menu item is selected, a window pops up immediately below that item. This window may be a dialog box, or contain a deeper menu level which in turn may activate other windows, etc. This approach allows the user to maintain a clear overview of the current UI status. Moreover, easy switching between various menu windows is possible. Although convenient, a mouse is not absolutely necessary to control this type of interface. Pull-down menus are frequently applied today in applications like word-processors and compilers/debuggers.

In addition to the already mentioned general requirements, the UI of the tele-education system should fulfil the following specific requirements:

**A: Adapted to the specific circumstances of a communication session**

When a remote lecture is in progress, commands should be effectuated by a minimal number of actions (keystrokes or mouse clicks) without loss of the user-friendliness.

**B: Maximum shared-screen area**

During the communication session, the UI control areas on the screen (menu bars, and windows) should be minimized to allow the shared-screen space to be as large as possible.

**C: Support of new or incidental users**

The UI should be designed to support incidental use like guest lecturing by minimizing the number of actions for low-level work with the tele-education system. E.g., a series of overhead sheets may be transferred using the system's document camera, requiring just
one keystroke for each sheet. This approach avoids the necessity for such users to train an extended set of editing and file commands.

To fulfil these requirements the UI specifications should include:
A: A variant of the pull-down menu method. A horizontal menu-bar only takes about 5 percent of the screen area. A drawback of the pull-down method is that it activates a menu-window over the work area as soon as a menu-bar item is activated. Therefore we prefer the following modification: the current menu level is displayed at the menu bar at the top of the screen. Mouse actions or keystrokes activate menu items. These could include deeper or higher menu levels, but the number of levels should be limited to three (for a few items to four). If necessary, windows should be used to obtain or present additional information, but should never be visible during execution of editing commands.
B: Shortcuts. To offer the more experienced user the possibility to switch rapidly between items of the various sub-menus, shortcuts should be defined (function keys) for the most frequently used commands.
C: An icon at the menu-bar should indicate the current menu status
D: The menu is to be mouse controlled, but working with only a keyboard is supported.
E: Status information (colour, thickness, font type, size, presentation-name, macro-library name, etc.) should be displayed at the menu-bar if applicable.
F: A 'help' window should be popped up by one of the function keys, providing more detailed user information on the current menu item.

Fig. 3.3 gives an impression of the resulting screen layout, including the three possible video window sizes.

![Diagram of screen layout](image)

**Fig. 3.3 Proposed VGA screen layout, including the three options for reduced-video windows**
3.2.2.7 Performance aspects

Taking into account the minimal terminal configuration (a 286 PC with 1 Mbyte memory) and the design-driving 9600 bit/s link, the system should be realized such that low performance never violates the effectiveness of the teaching process. In this context three classes of performance issues are distinguished: user interface response times, image display build-up times, and end-to-end transfer times of objects, pages, and video images (Issues such as resolution, colours and audio quality have been addressed in section 3.2.2.1).

The following performance is therefore required for:

A: UI response times
All actions via the UI, including editing, should be effectuated fast enough (within 1 second, Schoeffler [87]). Typing of text should be possible at nominal typing speed (10 characters/s).

B: Image display build-up
If the screen contents is to be redrawn completely, this action should be completed within a few seconds (preferably < 2 s). However, in the case of complex pages with a lot of handwriting information, a somewhat lower speed is allowed for object-oriented image build-up, which gives an illusion of a faster machine than with a line-by-line display mechanism, even where the latter is objectively faster.

C: End-to-end transfer times
Real-time tele-education requires on-line generated objects and retrieved pages to be transferred and displayed at the remote site immediately. Hence the effective transfer time should be < 2 seconds, corresponding to the requirement of the display build-up time. However, for reduced video a few seconds (but certainly less than 10 seconds) may be acceptable for the largest image size. Kenyon, White & Reid [85] suggest that 5 seconds would be a practical goal.

3.2.2.8 Utilities

Utility software should be available for hardcopy purposes, image file conversion, system diagnostics, and configuration of the system. By a hardcopy utility the pages of a presentation may be printed on a matrix or a laser printer. If a colour-printer would not be available, colours should be transformed into appropriate dither patterns. Image file conversion software may be useful for inputting images represented in a standardized format
3.2 Requirements and constraints

(like TIFF, GIF, etc., see Graef [89]). System diagnostics software should be used to check if the necessary peripherals are installed correctly and to test the actual communication links. Finally, a system configuration set-up program should be used for the generation of a configuration file providing information on the main application on the peripherals installed, certain communication parameters and a number of adjustable system parameters.

3.3 System design considerations

The purpose of this section is to translate the presented requirements into a conceptual design. After considerations regarding telecommunication network issues, architectural concepts are derived both from a software and a hardware point of view. Finally, a number of critical issues are identified which require development of special techniques.

3.3.1 Networks and configurations

As already discussed, to design a system independently of the locally available telecommunication infrastructure, the application-specific functions of the system should be separated from the lower network-oriented layers. However this ideal approach cannot be followed entirely in all situations concerned. Therefore we will first draw attention to various applicable networking concepts, discussing the impact on the application layer functionality of the envisaged terminal.

To realize a tele-education network two main options are distinguished: 1) use the public telecommunication infrastructure; 2) implement a dedicated (privately owned) network. In the former case, it is assumed that (only) the local Public Switched Telephone Network (PSTN) and leased lines would be available providing analog circuits. In the latter case, often some satellite communication means would be applied in view of the distances between the participating islands.

a) PTT infrastructure (analog)

The use of the PTT facilities allows dial-up or leased lines. In general, the latter may be conditioned more easily to enable bitrates up to 9600 bit/s. Audio handling can be realized using well-known bridge techniques either supplied as a general PTT service
('teleconferencing service') or as an extension of the university's PBX (Private Branch Exchange) facilities. Preferably automatic multiple call set-up should be supported. The teacher should be able to separately control the volume of the sound originating in the remote classes. For data communications, a multipoint network should be established supporting full-duplex links with bitrates of 9600 bit/s. If this rate cannot be achieved on a particular link, automatic fall-back procedures should be invoked implying lower feedback rates (inhibiting reduced video in the direction from the remote class towards the teacher) or even a lower forward bitrate. Hence a number of specific requirements for this situation apply:

A: Automatic multiple call set-up
This includes audio and data connections. The teacher should be provided with terminal and link status data.

B: Floor control issues
Provisions are preferred to control the audio-signal levels.

C: Data flow
All secondary stations will receive identical data from the primary station. The bitrate of this data stream should be automatically adapted corresponding to the capacity of the weakest link. Proper flow control mechanisms should be incorporated to avoid data overflow due to differences in effective link capacities, in CPU performances, or by application-specific effects.

D: Data link protocols
An adequate data link protocol should be applied to guarantee a reliable connection with sufficient BER (Bit Error Rate) level and to support link management functions. Proper mechanisms for rate fall-back should be applied to adapt gracefully to link degradation.

E: Feedback data
The data feedback signals from all the remote sites should be monitored continuously by the primary station, to react adequately on attention requests and to enable immediate video image switching at the teacher's display.

Fig. 3.4 depicts the corresponding network configuration. Obviously, instead of using two PSTN circuits, the data-signal may be relayed via a Public Data Network (PDN).

b) Dedicated satellite network

The obvious case is the installation of a satellite terminal at each of the tele-education facility sites, in order to be independent of the public telecommunications infrastructure. One of the attractive options is a VSAT (Very Small Aperture Terminal) network. VSAT
networks can be classified into three general categories: broadcast networks, point-to-point networks, and two-way interactive networks (Rana et al. [90]). In a broadcast network, a centralized hub station broadcasts packetized data, program quality audio, broadcast video, or a combination to a group of or all remote receive-only VSATs. Point-to-point networks provide voice, data, and image transmission between two locations without the requirement of a large hub earth station. A variant is a star network, whereby point-to-point circuits are provided from a centralized location to multiple remote locations. These networks usually employ Single-Channel-Per-Carrier (SCPC) transmission supporting a variety of applications including remote batch, file transfer, high-volume printing, and digital facsimile.

Two-way interactive networks are by far the most applied category. These networks offer a wide range of voice, video, and data services from a central hub station to a large number of remote VSATs in a star topology. The outbound transmission from the hub is broadcast to all VSATs over a high-speed time-division multiplexed (TDM) channel. By use of addressing in the TDM frame, a VSAT accepts only the messages directed to it. Inbound transmissions from the VSAT to the hub are carried over one or more channels using a dynamic multiple access protocol, so that many VSATs can share a satellite channel in low-activity applications.

A two-way interactive VSAT network including a central hub station will not be a cost-effective solution with a relatively small number of terminals. In that case, a more attractive approach would be to broadcast the audio and data signals of the primary station to all remote sites. This implies uplinking from a local PTT earth station, requiring a data
link from the university using, e.g., digital microwave LOS. Alternatively, at the primary location a dedicated private uplink station could be installed. In either case, all secondary stations could be equipped with a Receive-Only terminal. An uplink with, e.g., an SCPC channel with data rate 56 kbit/s, could be used to carry the multiplexed digital audio and data signals. In this context, Mneney [83] proposed a similar solution to realize a one-way tele-education service aimed at rural areas in African countries.

Feedback from the secondary stations is possible via terrestrial networks using the scribophone technique (Bordewijk [78]) which employs a single PSTN circuit for both voice and data. The latter may support handwriting, text-input and pointing. A phone-in configuration may be adopted, to reduce telecommunication costs. Note that only one satellite channel is in use, improving cost-effectiveness. The low investment costs and the relatively simple installation and operation procedures at the remote stations allow for flexibility and extendability with respect to the serviced target groups. Fig. 3.5 illustrates this satellite-based network concept.

This type of solution imposes the following specific requirements:

A: Voice-data multiplexing
At the primary station the mixed audio-signals should be digitized, compressed (if applicable) and multiplexed together with the outgoing data. Consequently, at the reception sites demultiplexing, decompression and digital-to-analog conversion is necessary.

B: Phone-in services
Feedback should be realized by a phone-in facility. Although audio-only might provide useful interaction, the inclusion of a data signal is preferred to allow inputs from keyboard, mouse and writing tablet. At the primary station, adequate provisions should be available to signal, select and accept the incoming phone-in requests.

C: Dataflow
Since flow control cannot be applied in a pure data-broadcast situation, the data rates should be adapted to the lowest performant station in the network.

D: Error correction
An appropriate Forward Error Correction (FEC) method should be applied on the satellite link to minimize the effect of bit errors in transmission of data. This requirement is omitted in the case of PCM transmission of the audio component.

In this context it is worthwhile to mention the highly attractive datacasting capabilities of the European D-MAC television standard. This allows narrowcasting of various data and digital audio signals, piggy-backed on a satellite broadcast TV signal, to remote sites equipped with appropriate low-cost PC-based receive terminals (Dobbie et al.
Fig. 3.5 A satellite-based network concept using a single broadcast channel and phone-in feedback

[89], Larock & Buelens [89]). The above mentioned tele-education concept including phone-in feedback could realize cost-effective distance teaching services directed to areas such as Western Europe, using the family of MAC standards.

Generally, the satellite network option requires quite sophisticated equipment, which is rather contradictory to the comments in paragraph 3.2.1 on limiting conditions and constraints. Therefore the terrestrial network option should be preferred, at least as a first stage, enabling small-scale pilot experiments.

3.3.2 Conceptual system architecture

The block diagram in Fig. 3.6 depicts a prototype terminal architecture seen from the hardware point of view, based on a 286, 386 or 486 (Intel) processor. Signal input devices comprise keyboard (for control, text and graphics), pointing device (control, pointing and marking), an electronic writing tablet (for handwritten graphics and annotations), a video-camera, a document scanner and a microphone. The display unit (i.e.,
VGA, with 16 or 256 colours) shows the shared-workspace window, the video window, and system control items. Communication can be established by a two-way data-link and a two-way audio-link. The use of analog or digital communication, applicable bitrates, and multiplexing issues, depend on the available telecommunication infrastructure.

The terminal software architecture should be modular to ensure that the addition of new media or communication interfaces, changes in the user interface or in the screen resolution, or other modifications, would have limited effect on the software. As indicated in Fig. 3.7, six main groups of software modules may be distinguished: control, display interface, communications, file & macro management, multimedia processing and user interface. The main function of the control group is to provide a non-preemptive task scheduler for the various communicating processes. The second group of modules provides the display management and driver functions, including windows control. Communications software resides in a third group of modules consisting of application-level, send and receive processes plus a multipoint communication control layer. All user actions are interpreted by the user interface module, which routes appropriate calls to other relevant modules. The group of multimedia processing modules include the handling of reduced video and the input and editing of text, structured graphics, raster-scanned
Fig. 3.7 Software architecture, showing the six groups of modules and the major data flows

graphics (document camera), and handwriting. Further the current virtual shared-screen buffer is controlled by these modules. The last group of modules encompasses presentation file management and macro-file management functions.

3.3.3 Critical technologies

The further development of a tele-education system as specified in the previous sections is now a rather straightforward process, except for a number of critical issues. Thus it is of vital importance that during an on-line lecture the shared screens remain synchronized with each other, in all circumstances. Delays of more than a few seconds between (alterations of) the teacher's screen and those of the remote classes would severely hinder the educational process. Unfortunately, video images and complex pages (pictures, complicated diagrams or drawings) tend to suffer transmission delays of several, up to tens of, seconds at 9600 bit/s, even where standard compression techniques are applied.
Therefore the next two chapters are dedicated to the development of an effective and efficient codec for the transmission of reduced video, designed for software-based implementation. Further, Chapter 6 proposes a number of practical improvements aimed at the use and processing of handwriting data in a real-time tele-education environment. Chapter 7 considers application protocols designed to reduce end-to-end page transfer times. Finally, Chapter 8 describes the overall integration of a PC-based tele-education terminal based on the requirements presented in this chapter, applying the delay-reduction techniques of the next chapters.

3.4 Summary and conclusions

Topographical barriers often prevent certain communities to benefit from the educational facilities of regular universities and schools. This chapter studied specific situations in which real-time tele-education can offer an effective solution to this problem. The University of the Netherlands Antilles (UNA) has been used as a case representing a broader range of similar situations, in order to identify the problem areas and to analyse various broad approaches to a solution. For the UNA, it has been concluded that real-time tele-education offers good prospects to extend effective and successful educational services to the remote communities.

This method is the closest replicant of regular face-to-face teaching. Only minor adaptions of teaching methods and curricula are required, with the advantages of comparatively low operational costs and a rather low psychological threshold to get such a method introduced. In addition to two-way audio communication, a visual component is imperative in real-time tele-education. Given the lack of wideband telecommunication channels to support the application of Instructional Television (ITV), the use of narrow-band visual signals should be applied. Therefore a requirement analysis has been performed aimed at the development of a multimedia PC-based system including audio, graphics and reduced-video services.

The impact of the local telecommunication infrastructure on such a system has been assessed, and two broad network concepts have been identified: a terrestrial multipoint network employing the Public Switched Telephone Network (PSTN) and a satellite-based concept for broadcasting the lecture to Receive-Only remote stations, with the possibility of phone-in feedback via the PSTN. A conceptual system architecture was discussed and a number of critical technologies identified, to be considered in the next chapters.
References in Chapter 3


Algra, T. [88], "Towards a tele-education network on the Netherlands Antilles and Aruba", University of the Netherlands Antilles, UNA-Cahier, No.27, (in Dutch), Curacao, 1988

Arnbak, J.C. [88], "Trans-border communications networks: technology-driven or social innovation?", European Management Journal, Vol.6, No.4, pp.330-337


Bates, A.W. [88], "Technology for distance education: a 10-year prospective", Open Learning, pp.3-12, November 1988


Borenstein, N.S. [91], "Multimedia Electronic Mail: will the dream become a reality?", Communications of the ACM, Vol.34, No.4, pp.117-119, 1991

Bruce, R., S. Elrod [92], "The electronic whiteboard", Byte, pp.166-167, July 1992

Byte [92], "Windows sells hardware", Byte, Vol.17, No.11, pp.20, 1992


Comm. N.A. Government [83], Report Committee to study the desirability of the establishment of an Open University of the Netherlands Antilles, Government of the Netherlands Antilles: Willemstad, Curacao, (in Dutch) 1983


Kenyon, N.D., T.A.White, G.M.Reid [85], "Behavioral and user needs for teleconferencing", *Proceedings IEEE*, Vol.73, No.4, pp.689-699, 1985

Lalor, G.C., C.Marrett [86], *Report on the University of the West Indies Distance Teaching experiment*, Kingston: University of the West Indies, 1986


Mneney, S.H. [83], "Audio-visual broadcast to rural areas over narrowband satellite channels", *Proc. IEEE Africon'83*, Nairobi, 1983


Rosenthal, L. [87], *The Eastern Caribbean Center Telecommunications Project*, University of the Virgin Islands, February 1987

4 LOW FRAME-RATE VIDEO CODING
FOR NARROWBAND TELE-EDUCATION

4.1 Introduction

This chapter presents a Low Frame-Rate Video Coding (LFVC) method, which was
developed as a subsystem for a real-time tele-education terminal. It encompasses a fast and
efficient coding scheme, appropriate for software implementation. Besides tele-education
applications, the method is useful for fast retrieval of video-sequences from multimedia
databases in PC networks. A summary version of this chapter was published in Algra [92b].

4.1.1 Multimedia Tele-education Terminal

Modern telematic technology enables the extension of the educational facilities of Universities
to remote locations, with real-time tele-education systems. Application in third-world
countries usually imposes a number of specific conditions and constraints. Firstly, the
available transmission bandwidth is often limited (e.g., 3.1 kHz analog telephone lines).
Secondly, to guarantee sufficient serviceability, standard equipment, without specialized
hardware has to be used. Thirdly, it should be possible to use the system without the need
to change the existing teaching methods and curricula, in order to keep the thresholds for
system introduction as low as possible: The teacher should be able to communicate with his
students in the same way as with face-to-face education. Hence, the tele-education system
should provide two-way audio communication, a common presentation screen (or work area)
- the replacement of the blackboard - and a visual component. The latter should enable the
students to watch the teacher (e.g., a small format head-shoulder image or a full screen
image during a demonstration session) and vice versa. The development of such a video
service is the subject of this chapter.
The block diagram in Fig. 4.1 (developed in the previous chapter) depicts a possible terminal configuration based on a 286, 386 or 486 microcomputer. Signal input devices comprise keyboard (for control, text and graphics), pointing device (control, pointing and marking), electronic writing tablet (for handwritten graphics and annotations), a video-camera, a document scanner and a microphone. The display unit (e.g., VGA, with 256 colours) shows the shared work space window, the video window and system control items (menubars, pop-up menus, messages, status information). Communication can be established through a two-way data-link and a two-way audio-link. The choice between analog or digital communication, bitrates, and multiplexing issues, depends on the local telecommunications infrastructure. With respect to telecommunication, this chapter deals only with the impact of the video transmission on the upper three levels of the OSI reference model (MacKinnon, McCrum & Sheppard [90]). In other words, it is supposed that an end-to-end link with sufficient capacity and reliability is available.

Recently, growing interest in multimedia conferencing technology has resulted in a large variety of publications. However, most systems which include video communication were developed for bitrates ≥ 64 kbit/s and require dedicated image (de)compression hardware (Watanabe, Kuroda & Hashimoto [87], Takizawa et al. [87], Yamaguchi, Wada & Yamamoto [86], Tanake et al. [86]). The latter also applies to the low bitrate codec described in Kelly [82]. The interactive conferencing application of Robinson et al. [91]
applies freeze-frame video, based on a software implemented predictive pyramid coding method, however, with only a moderate compression ratio.

In contrast, this chapter describes the development of a fast and efficient software-based low frame-rate video codec, which employs the interframe correlation of image sequences and obtains acceptable end-to-end transfer times at narrow bandwidths.

### 4.1.2 Low Frame-Rate Video Coding (LFVC) requirements

The development of the LFVC facility has been based on the following requirements, as partly formulated in Chapter 3:

a) *End-to-end transfer times as low as possible (certainly not more than 10 seconds).* This includes the time from video-frame digitization to display completion at the receiving terminal, for the range of bitrates discussed in i).

b) *Full duplex capacity.* Two LFVC signal streams should be handled simultaneously in opposite directions, without performance degradation.

c) *Three screen formats:* large (512x384 pixels, i.e., 80% of VGA full screen dimensions), medium (256x192 pixels), small (128x96 pixels). The small format is used to display the head-shoulder image of the teacher or, in opposite direction, of the auditorium or classroom during the presentation, and is inserted at a corner of the screen, with a short update-period. The larger formats provide better resolution, as needed for repeated updates of, e.g., the image of a blackboard, a demonstration set-up, or a closer inspection of the classroom.

d) The *image quality* should be equal to at least standard television quality.

e) *Extensibility to colour:* although the initial LFVC facility will handle black & white (continuous tone) images, extensibility to colour images should remain an option during algorithm and system development.

f) *Low cost system:* the system is aimed at small universities and institutes, and/or individual students. This leads to the requirement of a microcomputer-based terminal, with commonly available peripheral hardware.

g) *Serviceability:* with respect to the considered user groups (see point f), including universities in third-world countries, sufficient serviceability excludes the design and use of dedicated hardware. Therefore the complete LFVC system should be implemented as a software application for a microcomputer (PC).

h) *Processing requirements:* the tele-education terminal should be able to execute a variety of concurrent tasks like communication, input handling, file management. Therefore only 30% of the CPU processing capacity should be allocated to LFVC. Memory
occupation by LFVC functions should be < 512 kByte.

i) **Bandwidth (bitrate):** the data capacity allocated to the LFVC signal is 9600 bit/s, which assures operation when only analog telephone lines with advanced modems are available. However, the system should offer flexibility with respect to the available bandwidth and be able to operate over a wide range of bitrates (1200 to 56000 bit/s). Requirement a) should be met for bitrates equal to or greater than 9600 bit/s.

### 4.1.3 Image compression

From this list of requirements, it is clear that the considered source resembles a sequence of captured black & white video frames. It is assumed that normally the camera is fixed. Furthermore, the framegrabber employed is supposed to convert the camera signal into images of at least 512 by 384 pixels.

Assuming that each pixel is represented by 8 bits, 9600 bit/s transmission of an image of the largest format lasts at least 164 seconds. The amount of image data should be reduced by a factor 16 prior to transmission, in order to meet the maximum transfer time requirement. Digital images, in their canonical source representation, generally contain a significant amount of redundancy. Image compression aims at taking advantage of this redundancy to reduce the number of bits required to represent an image. Several excellent review articles on image compression have been published in the literature (e.g., Jayant [74], Jain [81], Netravali & Limb [80], Kunt *et al.* [85], Mussmann *et al.* [85]). A few books are wholly devoted to the subject (e.g., Rosenfield & Kak [76], Jayant & Noll [84], Netravali & Haskell [88], Rabbani & Jones [91]).

Three broad types of redundancy in digital images can be identified:

1) **Spatial redundancy**, which is due to correlation (or dependence) between neighbouring pixel values,
2) **Spectral redundancy**, which is due to correlation between different colour planes (e.g., in an RGB colour image) or spectral bands (e.g., aerial photographs in remote sensing)
3) **Temporal redundancy**, which is due to correlation between different frames in a sequence of images.

There are many approaches to image compression, but they can be categorized into two fundamental groups: **lossless** and **lossy**. In lossless compression (also known as **reversible** compression), the reconstructed image after compression is numerically identical to the original image on a pixel-by-pixel basis. Obviously, lossless compression is ideally desired since no information is compromised. However, only a modest amount of compression is
possible. In lossy compression (also known as irreversible compression), the reconstructed image contains degradations relative to the original. Note that these degradations may or may not be visually apparent. Much higher compression can be achieved than with lossless compression.

In the field of digital image compression, several standards are emerging. With respect to continuous-tone, still-frame, monochrome and colour images the Joint Photographic Experts Group (JPEG) standard was established, consisting of three elements: 1) a baseline system providing a simple and effective algorithm adequate for most image coding applications, 2) a set of extended system features such as progressive build-up allowing the baseline system to satisfy a broader range of applications and 3) an independent lossless method for the critical applications requiring that type of compression (Wallace [91]). With respect to sequential, continuous tone images, CCITT has standardized a coding algorithm (H.261) for video telephony and video conferencing at bitrates ranging from 64 to 1920 kbit/s (Liou [91], Plompens [90]). For storage and retrieval of moving images and sound using digital storage media with a combined rate of 1.0-1.5 Mbit/s, the Moving Pictures Expert Group (MPEG) standards are in the process of being finalized (Le Gall [91]).

The analysis as presented in the Sections 4.2 and 4.3 of this chapter will show that these standard methods are not suitable for the aimed LFVC application. Section 4.2 reveals the need for a computationally more efficient algorithm, based on reduction of spatial and temporal redundancy. Section 4.3 describes such an algorithm, based on intraframe as well as interframe coding. Section 4.4 reports implementation results and experiments, based on the derived algorithm.

4.2 Influence of coder complexity and compression factor on image transfer time

Much research effort has been spent on efficient image coding algorithms. Generally, the aim has been to maximize the compression ratio, under the constraints of a small or negligible loss of image quality and a complexity making hardware realisation (VLSI) feasible. One of the striking results of this international work is the abovementioned JPEG standard for the compression of still images. JPEG algorithms apply DCT with adaptive quantizing and subsequent entropy coding (Wallace [91]). For moving images, additional compression can be achieved by motion compensation (Plompens [90], Le Gall [91], Liou [91]).
These methods require a large computing capacity, especially when processing speed is an important requirement. Consequently dedicated (de)compression hardware is necessary to obtain acceptable processing power. In the literature, much less attention is paid to the systematic comparison of software-based algorithms with respect to speed and compression ratio, for a loss of image quality matched to the limited number of chrominance and luminance levels of standard PC screens such as VGA. An analysis of this kind, aimed at low frame-rate sequences ( < 0.5 frames/sec), will be offered in these sections.

The degree of complexity of a compression method is determined by the memory occupation ($M$, bytes) and the average number of instructions per pixel ($I$, ins/pel). Obviously, for a given algorithm a trade-off between $M$ and $I$ exists. The average number of bits per pixel of the compressed image representation will be referred to as the coding rate $R$. In order to be able to select suitable algorithms, we will first analyze the interrelated constraints on $R$, $I$ and $M$ which can be derived from the requirements stated. We will assume a symmetrical system, meaning that the CPU simultaneously executes a compress/transmit process and a receive/decompress process. We will derive the minimal achievable image transfer time given the maximal allowable CPU-load, as a function of the image size and the available transmission bandwidth.

The image transfer time (for short in the following: transfer time) is defined as the time interval between the frame-capture at the sending terminal and the completion of the image display process at the receiving terminal. The following processes contribute to this transfer time: image data acquisition into the PC memory, image compression (source coding), data formatting (including FEC coding), transmission, de-formatting, decompression, and display.

**Definitions (see Fig. 4.2):**

$T_{grab}$ = frame-grab time: the time required for the actual grabbing action and the input of the image data into the PC memory in the applicable format

$T_c$ = source-coding time: time needed to compress a complete image

$T_f$ = formatting time: time needed to format the compressed data, including generation and insertion of error correction codes if applicable

$T_t$ = transmission time

$T_{df}$ = deformatting time: time needed for deformatting including the processing of error correction codes if applicable

$T_d$ = decoding time: time needed to decompress the complete image

$T_{dis}$ = display time
The image transfer time $T_{if}$ is a function of these parameters:

$$T_{if} = f(T_{grab}, T_c, T_p, T_f, T_d, T_{df}, T_{dis})$$  \hspace{1cm} (4.1)

If all processes corresponding to $T_{grab}, T_c, T_p, T_f, T_d, T_{df}$ and $T_{dis}$ are executed serially, $T_{if}$ would be:

$$T_{if} = T_{grab} + T_c + T_f + T_t + T_{df} + T_d + T_{dis}$$  \hspace{1cm} (4.2)

$T_{grab}$ and $T_{dis}$ do not depend on the codec implementation, but they are determined by the available hardware and associated driver software (the frame-grabber and the display-adapter, respectively). Moreover, in practice $T_{grab}$ and $T_{dis}$ usually will have a duration less than 1 second. Therefore, these two parameters have a minor influence on $T_{if}$ and will be neglected further on. Formatting and deformatting times are assumed to be negligible.

Hence:

$$T_{if} = T_c + T_f + T_d$$  \hspace{1cm} (4.3)

Furthermore, symmetrical coding algorithms are assumed, meaning

$$T_c = T_d$$  \hspace{1cm} (4.4)

The minimal transmission time of one image is equal to
Fig. 4.3 Parallelisation of coding, transmission and decoding

\[ T_i = \frac{LR}{B} \]  \hspace{1cm} (4.5)

with \( L \) the number of pixels per image and \( B \) the channel data rate (bit/s). The coding time of one complete image is equal to

\[ T_c = L \frac{I}{c_l} \]  \hspace{1cm} (4.6)

with \( i_t \) the average duration of a CPU instruction and \( c_l \) the CPU-load of the coding process (i.e., the fraction of the CPU processing capacity allocated to the coding process). To speed up image transfer, the coding and decoding processes can be executed in parallel with the actual transmission process. This can be realized by partition of the image data into units of \( L_p \) pixels. As soon as a unit of \( L_p \) pixels is encoded, transmission of these data starts. Similarly, on the reception side the decoding process starts once a unit of data has been received. Fig. 4.3 visualizes this approach.

It is easily verified that \( T_i \) is equal to (see also Fig. 4.3):
4.2 Influence of coder complexity and compression factor

![Image](image.png)

**Fig. 4.4** Image transfer time $T_{if}$ as function of number of instructions per pixel, $I$, with coding rate $R$ as parameter, for $L = 512 \times 384$ pixels/image, and $B = 9600$ bit/s channel rate

\[
T_{if} = \begin{cases} 
2 L_p I \frac{i_t}{c_l} + \frac{L R}{B} & \text{for } \frac{R}{B} > I \frac{i_t}{c_l} \\
(L + L_p) I \frac{i_t}{c_l} & \text{otherwise}
\end{cases} \quad (4.7)
\]

If $L_p \ll L$:

\[
T_{if} \approx \begin{cases} 
\frac{L R}{B} & \text{for } \frac{R}{B} > I \frac{i_t}{c_l} \\
L I \frac{i_t}{c_l} & \text{otherwise}
\end{cases} \quad (4.8)
\]

It should be noted that $i_t$ not only depends on the CPU clock speed, but also is influenced by the actual (de)coding software, since it depends on the mix of instructions used (e.g., a multiply instruction has a longer duration than a shift instruction). Fig. 4.4 depicts $T_{if}$ as a function of $I$ for various values of $R$ with $B = 9600$ bit/s, $i_t = 1.0 \mu\text{sec}$ (typical for 286-PC,
with 12 MHz clock frequency). \( L = 512 \times 384 \) pixels, and \( c_i = 0.15 \), which corresponds to a joint load of 30% by simultaneous coding and decoding.

The graph stresses the point that the selection of an optimum algorithm resulting in a small \( T_f \) not only depends on \( R \), but also on \( I \). For example, it is not useful to apply algorithms with better compression performance than 1 bit/pel, if the number of instructions per pixel \( I \) exceeds 26. In Rabbani & Jones [91] a review is given of image compression methods, including a useful comparison of image quality versus compression ratio. Also, complexity issues are assessed and the number of operations per pixel is estimated for most methods. In their analysis, Rabbani & Jones describe the complexity of the various algorithms by the average number of operations per pixel (i.e., multiplications, additions and comparisons). Here, these results are re-interpreted in terms of our analysis in order to get an initial impression of \( T_f \).

Table 4.1 lists the results for methods with \( R < 1 \) bit/pel, but acceptable to good (reconstructed) image quality. The JPEG DCT variant uses the ultra fast DCT version of Arai et al. [88]. The two vector-quantizing variants are mean residual tree structured vector quantizing (M/RTVQ) and interpolative residual tree structured vector quantizing (I/RTVQ), see Rabbani & Jones [91]. A subband coding algorithm (combined with VQ) is also included (Westerink [88], Rabbani & Jones [91]). More detailed information on specific algorithms is deferred to Section 4.3. For each method the average number of assembler instructions per pixel is estimated and listed in column 'I'. The resulting values are indicated in the Figures 4.4 and 4.5.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>operations/pixel</th>
<th>I (ins/pel)</th>
<th>R (bit/pel)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( x )</td>
<td>( +, - )</td>
<td>( &lt;, &gt; )</td>
</tr>
<tr>
<td>JPEG (DCT)</td>
<td>2.25</td>
<td>9.25</td>
<td>1</td>
</tr>
<tr>
<td>M/RTVQ</td>
<td>40</td>
<td>42</td>
<td>2.19</td>
</tr>
<tr>
<td>I/RTVQ</td>
<td>43</td>
<td>45</td>
<td>2.19</td>
</tr>
<tr>
<td>SBC/VQ</td>
<td>64</td>
<td>124</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.1 \( I \) (instructions/pixel) and \( R \) (bit/pixel) values of a selection of algorithms (as from Rabbani & Jones [91])

\(^1\) Feig [90] describes a DCT algorithm of comparable speed
4.2 Influence of coder complexity and compression factor

![Graph showing Tff (s) vs. I (ins/pel)](image)

Fig. 4.5 Image transfer time $T_{ff}$ as function of number of instructions per pixel, $I$, with coding rate $R$ as parameter, for $L = 512 \times 384$ pixels/image, and $B = 56$ kbit/s channel rate

Interpretation of these results leads to the following conclusions:

1) JPEG DCT gives the best performance of the standards in terms of $T_{ff}$
2) For all these cases, $T_{ff}$ is constrained by the value of $I$, while the seemingly good compression ratio (i.e., $R < 1$ bit/pel) does not contribute to a short transfer time.
3) For $B = 9600$ bit/s, the achieved $T_{ff}$ values do not comply with the requirements stated in 4.1.2.
4) Only reduction of $I$ would effectively improve the transfer speed.

Motion-compensation (MC) techniques can effectively reduce temporal redundancy in image sequences (Le Gall [91]). However, Conditional Replenishment (CR) (Candy et al. [71]) is an interframe coding method with significantly lower complexity. Each $N$-pixel block of the image is compared to the corresponding block of the previous image. The block is replenished only if the contents are changed due to motion (for a discussion on detection criteria and methods refer to Section 4.4).

Let $m$ be the fraction of changed (and freshly encoded) blocks, and $I_{CR}$ denote the number of instructions/pixel required by the CR process, then the average number of instructions per pixel $I'$ becomes:
Fig. 4.6 Image transfer time $T_{tf}$ with Conditional Replenishment as a function of the number of instructions per pixel, $I$ ($L = 512 \times 384$; $B = 9600$ bit/s; $m = \text{fraction of changed blocks per image}$).

$$I' = m \ I + I_{CR} \quad (4.9)$$

The effective coding rate is reduced to

$$R' = m \ R + \frac{1}{N} \quad (4.10)$$

The term $1/N$ in (4.10) is due to the addition of a CR status bit for each block. Substitution of $I$ and $R$ in (4.8) by $I'$ and $R'$ respectively gives:

$$T_{TF} = \begin{cases} \frac{L(mR + \frac{1}{N})}{B} & \text{for } R > \frac{I}{c_L} + \frac{I_{CR}}{m} + \frac{1}{m} \\ \frac{L(mI + I_{CR})}{c_L} & \text{otherwise} \end{cases} \quad (4.11)$$

Since it appears that $m \ll 1$ for video sequences with a static background (Candy et al. [71]), significant reduction can be realized for both $R$ and $I$, even at lower frame rates (see Fig. 4.6). The conclusion is that CR, in combination with an appropriate intra-frame coding method, gives the best results. The selection of such an intra-frame method is described in Section 4.3, which also gives a description of the CR method.
4.3 An effective algorithm for low frame-rate video coding

4.3.1 Introduction

This section first reviews image compression methods, and discusses issues which impact the subsequent determination of the best combination of algorithms and methods for LFVC, described in the paragraphs 4.3.3 and 4.3.4. Issues specifically related to motion video (frame-rates of 25 or 30 Hz) and colour coding are not discussed. The following paragraph deals with the basic operations truncation (number of grey levels) and entropy coding, and describes predictive coding, transform coding, vector quantizing, block truncation coding (BTC) and subband coding (SBC). Then a suitable intraframe compression method is derived, followed by the development of a conditional replenishment approach.

4.3.2 Review of image compression methods

Video signals usually are digitized into 8 bit/pel per colour, i.e. 24 bit/pel and 8 bit/pel for RGB signals and monochrome continuous tone images, respectively. If the number of bit/pel is decreased, contouring effects become apparent, especially in regions with minor luminance variations. Often, contouring effects are not visible for representations with \( \geq 5 \) bit/pel but become unacceptable if \( \leq 3 \) bit/pel (Huang et al. [67], Rosenfield & Kak [76]). Subjective evaluation tests confirm this perceptual property (Inoue, Yoroizawa & Okubo [84]). Contouring effects can be masked to a certain extent by dithering techniques, involving the addition of pseudo-noise to the video signal before A/D-conversion. In this way, even 3 bit/pel images show tolerable to moderate quality (Limb [69], Lippel, Kurland & March [71], Lippel & Kurland [71]).

4.3.2.1 Entropy coding

Reversible datacompression, often entitled entropy coding, aims at reduction of the redundancy of a stream of source-letters, produced by a source of which the alphabet exhibits a non-uniform probability distribution. The well-known and often applied code construction procedure of Huffman [52] produces optimal fixed-to-variable-length codes, i.e., to the various source letters codewords of different length are assigned, resulting in the lowest possible coding rate. The resulting average number of bits per letter approaches the entropy
of the source

\[ H = - \sum_{i=1}^{N} p_i \log_2 p_i \]  

(4.12)

where \( p_i \) denotes the probability of source letter \( i \).

Encoding of variable-length sequences of source outputs into codewords of constant length is called \textit{variable-to-fixed-length} coding. It can be considered a generalization of run-length encoding and is a technique that seems especially attractive for a skew source (i.e., where the frequency of some output letters very much exceeds that of others). Variable-to-fixed-length coding was investigated by Tunstall [67], who described an optimal code construction procedure. Jellinek and Schneider [72] proved that for every source a code can be constructed that approaches the source entropy. Ziv [90] proved that for \( k \)'th order Markov sources, a variable-to-fixed-length code performs better than the best fixed-to-variable code. Liu [92] proposed a practical application of variable-to-fixed-length coding with differential chain coding of line drawings. In Algra [92a] efficient variable-to-fixed-length algorithms are compared to Huffman coding and a modified version of Tunstall's coding (MTC) algorithm is proposed, with improved performance. A detailed description of MTC can be found in Chapter 5.

Often, entropy coding is the last part of the source-coding chain in an image encoder. Thus, the JPEG image compression standard applies Huffman coding to the DCT coefficients (Wallace [91], Rabbani & Jones [91]). The same principle is used in the visual telephone standard H.261 for image compression (Plompem [90]).

4.3.2.2 Predictive coding

a) \textit{Differential Pulse Code Modulation (DPCM)}

In a general predictive coding scheme (Makhoul [75]), the correlation between the neighbouring pixel values is used to form a prediction for each pixel. By far the most common approach to predictive coding is \textit{differential pulse code modulation} (Jayant & Noll [84], Schilling, Scheinberg & Garodnick [78]). In a lossy DPCM scheme, \( m \) pixels within a 'causal' neighbourhood of the current pixel \( x_k \) are used to make a linear prediction \( x_k' \) of its value, namely

\[ x_k' = \sum_{i=0}^{m-1} \alpha_i x_i \]  

(4.13)

where the \( \alpha_i \)'s are the prediction coefficients. The differential (error) image, \( e_k \), is
constructed as the difference between the prediction and the actual value; i.e.,

\[ e_k = x_k - x'_k \]  \hspace{1cm} (4.14)

In order to reduce the bitrate, the differential image in lossy DPCM is quantized prior to encoding and transmission. A block diagram of a DPCM coder and decoder system is shown in Fig. 4.7, where \( e_k^* \) represents the quantized differential image. In order to avoid accumulation of quantization errors in the decoder, the prediction at the coder side is also based on the reconstructed values. This is accomplished by containing the quantizer within the prediction loop as shown in the coder diagram of Fig. 4.7.

If the sum of the prediction coefficients (eqn. 4.13) is made slightly less than one (leaky prediction), the effects of channel errors is reduced. The following are typical examples of linear predictors (for location of the neighbouring pixels refer to Fig. 4.8):

\[
\begin{align*}
x' &= 0.97 A & \text{1st-order, one-dimensional (1-D) leaky predictor} \\
x' &= 0.50 A + 0.50 C & \text{2nd-order, two-dimensional (2-D) non-leaky predictor} \\
x' &= 0.75 A - 0.50 B + 0.75 C & \text{3rd-order, 2-D non-leaky predictor} \\
x' &= A - B + C & \text{3rd-order, 2-D non-leaky predictor} \\
x' &= 0.90 A - 0.81 B + 0.90 C & \text{3rd-order, 2-D leaky predictor}
\end{align*}
\]
A substantial portion of the compression achieved by a lossy DPCM scheme is obtained by the quantization of the differential image. A quantizer is defined by a finite set \( \{ r_i, i=0,\ldots,N-1 \} \) referred to as reconstruction levels and a set of points \( \{ d_i, i=0,\ldots,N \} \), referred to as decision levels. The quantizing operation effectuates that an input value \( e_k \) on the interval \( (d_i,d_{i+1}] \) is mapped to the discrete value \( r_i \). For systems encoding the quantizer output levels using an entropy coding method, the optimum quantizer is uniform if \( e \) is distributed in accordance with a Laplacian probability density function, i.e., the decision regions all have the same width (Wood [69], Rabbani & Jones [91]).

b) **Adaptive Differential Pulse Code Modulation (ADPCM)**

A major limitation of the DPCM principle considered so far is that both the predictor and the quantizer are fixed throughout the image. DPCM schemes can be made adaptive in terms of the predictor, the quantizer or both. Adaptive prediction attempts the continuous optimization of the predictor coefficients with respect to the current image statistics. The result is reduction of the prediction error prior to quantization. This in turn decreases the quantization error signal and gives better image quality. An example of a computationally efficient adaptive predictor is the two-dimensional third-order median predictor (Murakami et al. [87]), which avoids the need to transmit separately the varying prediction coefficients, since the predictor state calculation is based on previous (reconstructed) pixels only. The algorithm is defined as follows:

\[
\begin{align*}
    x'_k &= \text{Med} \{ A, C, 0.5A+0.25C+0.25D \} \\
    e_k &= x_k - x'_k
\end{align*}
\]  

(4.15)

where \( \text{Med}\{p,q,r\} \) denotes the median value of \( p, q \) and \( r \). Adaptive quantization schemes vary the number of reconstruction levels on a block-by-block basis. It is well known that the luminance sensitivity of the human visual system (HVS) decreases in the picture areas of high-contrast detail. In these areas, large quantization errors can be masked by the HVS.
(Netravali & Prasada [77]). In the literature various systems are proposed; most of them select one member from a set of available quantizers for each block of pixels. The selection criterion should be based on a HVS distortion measure. Instead of this computationally intensive method, simpler distortion measures are used commonly, such as mean squared error (MSE) or the sum of the absolute error. A second point to note is that the coded signal needs some overhead to inform the decoder about the selected quantizer.

c) Performance and complexity issues of DPCM and ADPCM
In general, received image quality is poor to moderate if \( R = 1 \) bit/pel (i.e., a quantizer with only two reconstruction levels, known as delta modulation) due to poor edge representation and pronounced edge busyness (Schilling et al. [78]). Acceptable to good quality can be achieved at 2 or 3 bit/pel. However the average coding rate can be made significantly lower if entropy coding is applied on the DPCM coder output. Except for adaptive quantization and high order prediction, the DPCM algorithms are computationally less intensive than the majority of alternative compression methods.

d) Interframe coders
Three-dimensional prediction can be applied in interframe coding (Haskell et al. [72], Haskell, Connor & Mounts [73], Candy et al. [71], Haskell & Schmidt [75]), where pixels are categorized into classes of stationary and moving regions. For each separate class an optimal set of sampling procedures, predictors, interpolation filters and quantizers is used. For instance, pixels deviating no more than 1.5\% of the maximum luminance amplitude from their temporal predecessor, can be considered as unchanged. Regions with a minor number of changed pixels are coded with reduced temporal resolution, while regions characterized by many changed pixels, corresponding to the faster moving objects in the image, are encoded with a reduced spatial resolution. The latter approach exploits the reduced sensitivity of the human visual system (HVS) for details with respect to moving objects.

e) Conditional Replenishment (CR)
This technique is developed by Mounts [69] and subsequently refined by Candy et al. [71], Haskell et al. [72], Limb & Pease [71], Pease & Limb [71], et al. CR is a variant of the above mentioned interframe coders. Only those pixels which deviate significantly from the corresponding pixel in the previous frame, are encoded (e.g., by PCM or DPCM) and transmitted. For visual telephone applications (i.e., camera in fixed position; head-shoulder image) acceptable results are reported with average coding rates of 1 bit per pixel. A remarkable observation is that the number of changed pixels remains less than 10\% most of
the time (Candy et al. [71]).

Obviously the encoded active pixels must be accompanied by address information, in order to enable the receiving terminal to determine their screen position. Because changing pixels tend to occur in clusters, it is usually more efficient to address these clusters. Furthermore, CR requires buffering to compensate temporary variations in the coding bitrate. The required buffer size depends on the statistical properties of the image sequence. The probability of buffer overflow can be reduced by means of spatial subsampling (i.e., a temporary reduction of the resolution), selective replenishment (a form of interlacing, enabling 2:1 temporal subsampling) or a temporary increase of the detection threshold.

f) Motion Compensation

In general it is a simplification if, in interframe prediction, the frame period is considered as the basis for the periodicity of the video signal: In moving images a pixel \(x(p,q,r)\) is not maximally correlated with pixel \(x(p,q,r-1)\) of the previous frame\(^2\), but dependant on the speed and the direction of the moving object - with another pixel \(x(p+\delta_p,q+\delta_q,r-1)\). Thus optimized interframe predictors must incorporate some motion compensation technique (Ishigura & Linuma [82]). For visual telephone applications, this results in reduced coding rates, down to 0.15 to 0.3 bit per pixel. Motion compensation techniques are applied e.g., in visual telephony (Plompen et al. [86], Plompen [90]) and High Definition Television (Haghiri & Vreeswijk [90]). The high complexity level of these algorithms requires hardware implementation to gain acceptable encoding and decoding times.

4.3.2.3 Transform Coding

Transform coding is sometimes termed Block Quantization, since it implies that the input \(x(n)\) is divided into blocks of \(N^K\) samples, with each block converted to a set of \(N^K\) coefficients by a matrix transform. The coefficient \(K\) is the dimension of the transform (1,2 or 3). Subsequently, the transform coefficients (e.g., Fourier coefficients) are quantized and encoded. The receiver reconstructs the original sample blocks by application of the inverse transformation. When each coefficient (or group of coefficients) is quantized by a quantizer which is optimized for that specific coefficient with respect to source statistics and perception effects, then significant compression can be achieved. Possible optimization techniques include optimal bit allocation and zonal sampling (Jayant & Noll [84]). The visibility of the

\(^2\) \(p\) and \(q\) denote the horizontal and vertical pixel coordinates, respectively; \(r\) denotes the frame index
effects caused by transmission bit errors is less than for predictive methods. Here, an erroneous coefficient will only effect a limited region, namely a block of \( N^K \) pixels and is smeared out over the block area.

Frequently applied transforms are the \textit{Discrete Cosine Transform} (DCT) and the \textit{Discrete Fourier Transform} (DFT). For image processing applications, the forward 2-D DCT of an \( N \) by \( N \) block of pixels is often defined as (Chen & Pratt [84]):

\[
F(u,v) = \frac{4C(u)C(v)}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i,j) \cos\left[ (2j+1) \frac{u\pi}{2N} \right] \cos\left[ (2i+1) \frac{v\pi}{2N} \right]
\]  

(4.16)

The inverse 2-D DCT is

\[
f(i,j) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v) F(u,v) \cos\left[ (2j+1) \frac{u\pi}{2N} \right] \cos\left[ (2i+1) \frac{v\pi}{2N} \right]
\]  

(4.17)

where

\[
C(w) = \begin{cases} 
\frac{1}{\sqrt{2}} & \text{for } w=0 \\
1 & \text{for } w=1,2,...,N-1 
\end{cases}
\]

(4.18)

Another interesting transform is the \textit{Discrete Walsh Hadamard Transform} (DWHT), having only 1 and 0 matrix coefficients (Andrews, Jane & Pratt [69]). Consequently, DWHT transforms can be calculated with only additions and subtractions. However, this reduction in complexity has to be paid with a lower SNR (as compared with DCT and DFT).

Although transform methods result in considerable compression ratios with good image quality when combined with a proper entropy coding method (e.g., the JPEG DCT), their drawback is the inherent computational complexity. Several well-known algorithms (Vetterli & Nussbaumer [84], Lee [84]) require 29 additions and 13 multiplications for an 8-point 1-D DCT. However, in the context of the JPEG DCT algorithm, a scaled version of the DCT can be used, further reducing the number of multiplications to 5 (Arai et al. [88]). Therefore an 8x8 2-D DCT can be computed as 16 1-D DCT’s of size 8, requiring 2.25 multiplications, 9.25 additions and 1 comparison per pixel (Rabbani & Jones [91]). This is similar to an estimated 27 ins/pel.

4.3.2.4 Subband coding (SBC)

In subband coding the signal spectrum is split into 2 or more subbands by means of a series of bandpass filters (Woods & O’Neill [86]). After conversion to lowpass, the filter outputs
are sampled at a frequency of twice the filter bandwidth. Finally, these sampled signals are encoded by PCM or DPCM. The advantage is that each subband signal can be encoded with an accuracy adapted to the best perceptual criteria for that particular part of the signal frequency spectrum. This process can be implemented two-dimensionally, in order to fully exploit the characteristics of images. High compression ratios combined with good image quality results are reported, especially when the subband outputs are vector quantized; better than 1 bit/pel (Westerink [89]). However, the applied filtering structure requires a significant amount of computation per pixel. Executing the image decomposition process into e.g., 16 bands with 32-tap filters, requires 64 multiplications and 124 additions per pixel (Rabbani & Jones [91]). The synthesis operation in the decoder takes even more operations per pixel.

4.3.2.5 Block Truncation Coding (BTC)

The objective of block truncation coding (Healy & Mitchell [81]) is to quantize the pixels of an $N \times N$ block in such a way as to preserve its mean and variance. A two-level quantizer is attractive because of its simplicity. The first step is to measure, quantize, and transmit the mean, $\mu$, and the variance, $\sigma^2$, of the block. Next, the pixels of the block are quantized to two levels using $\mu$ as threshold, i.e., a 1 or a 0 is sent depending on whether $x_{ij}$ is greater or less than $\mu$, respectively. At the receiver the first step is to evaluate $q$, the number of 1's received for the block. The number of zeros is then $p = N^2 - q$. The next step is to evaluate the two quantizer levels that preserve mean and variance. These are easily shown to be

$$L_0 = \mu - \sigma \sqrt{q/p} \quad L_1 = \mu + \sigma \sqrt{p/q} \quad (4.19)$$

Finally, each pixel is decoded into $L_0$ or $L_1$, depending on whether a 0 or a 1 was received, respectively. Using blocks of size 4x4 pixels, and 8 bits each for $\mu$ and $\sigma$, the coding rate
is 2 bit/pel. If μ and σ are coded jointly using 10 bits, the data rate is 1.625 bit/pel. Decoded BTC images typically reproduce sharp edges of large objects fairly well. Also random texture, although not produced with the same accuracy, often looks well enough to be acceptable in many applications. However, in low-detail areas, where brightness varies slowly, contouring and block-edge discontinuity due to coarse quantization are often visible. A computationally efficient variant is absolute moment block truncation coding (AMBTC) with similar results (Lema & Mitchell [84]). Rabbani & Jones [91] took stock of the number of operations per pixel for adaptive AMBTC and showed the highly efficient nature of this family of algorithms. However, the achievable coding rate is moderate (R > 1 bit/pel).

4.3.2.6 Vector Quantization (VQ)

A fundamental result of Shannon's rate-distortion theory is that better performance can always be achieved by encoding vectors instead of scalars, even if the data source is memoryless, e.g., consists of a sequence of independent random variables. While compression schemes as transform coding operate on vectors also, and achieve significant improvement over PCM, the quantization is still accomplished on scalars and hence, in a Shannon sense, these systems are inherently sub-optimal. Better theoretical performance is always achieved by coding vectors instead of scalars, even if the scalars have been produced by preprocessing the original input data in order to make them uncorrelated or independent (Gray [84]).

The basic principle of VQ is to map an N-dimensional vector x onto a N-dimensional vector y, by

\[ y = Q(x) \]  \hspace{1cm} (4.20)

where \( Q(.) \) is the vector quantization operator. The reconstruction vector \( y \) typically takes one of a finite set of vectors

\[ Y = \{ y_i, i = 1, \ldots, c \} \]  \hspace{1cm} (4.21)

The set \( Y \) is generally referred to as the codebook containing \( c \) vectors. The coding rate is then given by \( 1/N \log_2 c \) bit/pel. The vector quantization procedure can be described as follows. First, the \( N \)-dimensional vector \( x \) is constructed from the input signal. For image coding, blocks can be defined within that image consisting of \( N \) pixels. Next, the best fitting codebook vector \( y_i \) is found by using a minimum distortion or nearest neighbour selection rule:
\[ Q(x) = y_i \iff d(x, y_i) \leq d(x, y_j), \quad j = 1, \ldots, c \] (4.22)

Here \( d(x, y) \) is the distance or distortion measure between the vector \( x \) and \( y \).

![Diagram showing partitioning of two-dimensional space into cells. Each vector in cell \( C_i \) is quantized to code vector \( y_i = Q(x) \).](image)

Fig. 4.10 Partitioning of two-dimensional space into cells. Each vector \( x \) in \( C_i \) will be quantized into the code vector \( y_i = Q(x) \).

Fig. 4.10 illustrates the process, in which two-dimensional space is partitioned into cells. The shape of each cell is uniquely determined by the positions of the code vectors and by the distortion measure. All input vectors that lie within cell \( C_i \) are quantized with its codebook vector \( y_i \). The optimum coder/decoder combination for a given compression ratio contains a codebook minimizing the variance of the quantization error. For such a combination the next two properties can be proven (Gray [84]):

**Property 1:**
Given the goal of minimizing the average distortion and given a specific codebook \( Y \), no memoryless quantizer encoder can do better than select the code word \( Q(x) \) in \( Y \) that will yield the minimum possible distortion.

**Property 2:**
Given an encoder \( Q(x) \), no decoder can do better than that which assigns to each coded output vector \( y = Q(x) \) the generalized centroid of all input vectors encoded into \( Q(x) \).

The centroid of cell \( C_i \), denoted by

\[ y_i = \text{cent}(C_i), \] (4.23)
is that vector \( y \) which minimizes the distortion:

\[
D_i = \int_{x \in C_i} d(x, y)p(x)dx
\]  

(4.24)

where \( p(x) \) is the multidimensional probability density function of \( x \). However, in practice the pdf \( p(x) \) is usually unknown. Then the codebook can be designed, using a \textit{training set} of images, representative for the application. For this purpose the LBG-algorithm can be applied (Linde, Buzo & Gray [80]):

**LBG algorithm:**

- **Step 0** Given: A training set of images and an initial codebook.
- **Step 1** Encode the training set into a sequence of output vectors \( y \), using the given nearest neighbour rule. If the average distortion is small enough, quit.
- **Step 2** Calculate the new codebook vector \( y_i \) for each cell \( C_i \) by calculating the centroid of the training vectors that were classified to that cell in step 1. Go to step 1.

Such iterative improvement algorithms will not, in general, yield truly optimum codes. It is often useful, therefore, to enhance the algorithm’s potential by providing it with good initial codebooks (Gray [84]). The simplest method to generate such an initial codebook is to take the first \( c \) vectors of the training set.

VQ encoders are characterized by high complexity. The encoding process is computationally intensive, because for each input vector a full search through the complete codebook must be performed (number of operations \( O(Nc) \)). Therefore much research effort has been spent on the development of methods reducing the encoder complexity by decreasing the code-book size. Three general approaches to this problem are \textit{tree-structured codebooks}, \textit{product codes} and \textit{classified VQ}.

When a \textit{tree-structure} is imposed on the code-book (Gray [84]), with each node having \( m \) branches and with \( p = \log_m c \) levels to the tree, the computational cost is reduced to \( O(Nm\log_m c) \), since only certain branches of the tree are examined. A tree-structured codebook can never perform better than a single-level full search codebook, since a tree structure effectively limits the possible codevectors, once a particular branch has been selected.

The \textit{product code} approach involves a codebook that is formed as the Cartesian product of several smaller codebooks. If a vector can be decomposed into certain independent
features, a separate codebook can be developed to encode each feature. Several types of product code VQ's are described in literature: among others Gain/Shape VQ, Interpolative/Residual VQ, and Mean Separating VQ.

In a mean-separating VQ (Gray [84]), one first uses a scalar quantizer to code the sample mean of a vector (defined by $N^{-1} \sum x_i$). Then the coded sample mean is subtracted from all the components of the input vector to form a new vector with approximately zero sample mean. This new vector is then vector quantized (see Fig. 4.11). The basic motivation here is that in image coding the sample mean of pixel intensities in a small rectangular block represents a relatively slowly varying average background value of pixel intensity around which there are variations. In this way the codebook size can be reduced significantly, but at the cost of compression ratio, since MSVQ does not exploit the correlation between the mean and the difference vector.

In classified VQ (Ramamurthi & Gersho [86]) a number of different codebooks are developed, each designed to encode blocks of pixels that contain a specific type of feature, e.g., a horizontal edge, a vertical edge, completely uniform area, etc. The codebook used to encode a particular block is determined by a classifier capable of discriminating between the different types of features. The rationale for this approach is that numerous small codebooks, each tuned to a particular class of vectors, can provide comparable image quality with lower computational complexity than with a single large codebook.
4.3 An effective algorithm for low frame rate video coding

Fig. 4.12 Typical peak SNR for various compression methods as a function of coding rate (from Rabbani & Jones [91])

4.3.2.7 Comparison of different methods

Fig. 4.12 compares Peak signal-to-(reconstruction) noise error (PSNR) values of a number of compression methods, for a particular 512 x 512 image (Rabbani & Jones [91]). Note that the results of the DPCM schemes can be improved by additional entropy coding. None of these methods meets the complexity requirement of $I < 15$ ins/pel as derived in Section 4.2. This motivates further search for adequate compression methods for a tele-education system.
4.3.3 Fast and efficient LFVC intraframe compression

4.3.3.1 Initial selection

From the analysis of Section 4.2 the following requirements for a useful LFVC algorithm are postulated:

a) \( R < 1 \) bit/pel
b) \( I < 15 \) ins/pel
c) Application of Conditional Replenishment

Implementation requirements following from Section 4.1 are

a) 100% software implementation
b) memory occupation \(< 512\) Kbyte
c) three screen formats: 512x384, 256x192, 128x96 pixels
d) black & white image, with potential extensibility to colour
e) VGA monitor (640 x 480 pixels, 256 'colours')

Extensibility to colour LFVC in combination with a VGA-type screen (256 simultaneously displayed 'colours') leads to the allocation of 16 luminance ('grey') levels and 16 basic colours, resulting in acceptable image quality. In other words, the pixels of black & white LFVC should be reconstructed by 4 bit/pel, offering the possibility to base the digitization process also on 4 bit/pel. The advantages of this approach are:

a) Less memory occupation is needed for frame storage.
b) It enables computationally efficient algorithms by using look-up tables (LUT).
c) It enables application of Modified Tunstall Coding (MTC) due to the limited number of source letters. MTC is a computationally efficient method (see Chapter 5).

Points b) and c) will be shown to be the basis for ultra fast implementations in the following analyses. A drawback may be the introduction of extra quantization noise prior to the actual coding process, resulting in some reduction in compression effectiveness. In 4.3.2 it has been mentioned that representation by 4 bit/pel does not result in annoying contouring artefacts in most cases, and that, moreover, dithering techniques can soften these contours.

In order to select that intra-frame method in combination with CR which best meets
the criteria, we will follow a two-step approach: 1) Initial selection of a number of candidates, based on estimations of \( R \) and \( I \) derived from literature data; 2) Final selection, based on experimental comparison of the initial candidate algorithms.

The results of the survey of paragraph 4.3.2 revealed that the following methods should be excluded as candidates:

a) Block Truncation Coding (BTC)
The achievable coding rates in order to gain acceptable image quality are greater than 1 bit/pel.

b) Subbandcoding (SBC)
The bottleneck of SBC is its computational complexity, caused by the two-dimensional bandpass filter structure, which would result in high \( I \) values and consequently unacceptable transfer times.

c) Transform coding
Although transform coding appears to be more computationally efficient than SBC, even the fastest algorithms require a considerable amount of operations per pixel, resulting in moderate transfer times.

Of the remaining methods, a number of candidates and their variants should be considered, however with some remarks:

a) DPCM
In combination with an appropriate entropy coding method, the resulting coding rates will be \( \ll \) 2 bit/pel. The relatively low complexity of DPCM leads to acceptable values of \( I \). Therefore DPCM variants will be investigated, including 1 and 2-dimensional coding, various orders of prediction, and adaptive prediction.

b) Vector Quantization
VQ offers good compression ratios, but its inherently high level of complexity hinders application with LFVC. However, we will investigate if a version based on look-up tables could result in the required transfer times with acceptable image quality.

These experiments are described in the subsequent paragraphs. Two image sequences were selected, which can be considered as typical examples of tele-education scenes. The sequence *HEADS* consists of 30 small format frames (128 x 96 pixels) of a talking person in front of the fixed positioned camera (head-shoulder image). The sequence *BOARD* consists of 10 large format frames (512 x 384 pixels) of a teacher writing on a white-board, also with fixed positioned camera. For both sequences the frame rate was \( \approx 0.3 \) images/sec.
Fig. 4.13 Histogram of amplitude distribution of DPCM quantizer output (sequence BOARD)

4.3.3.2 Differential Pulse Code Modulation (DPCM)

a) First order, 1-dimensional DPCM (DPCM, 1st order, 1-D)

Encoder algorithm:

\[ e_k = x_k - x'_{k-1} \]
\[ x'_{k} = x'_{k-1} + e_k \]  \hspace{1cm} (4.25)

By application of a 16-level uniform quantizer, \( x_k \) will be coded without additional quantization error. Then (4.25) can be rewritten:

\[ d_k = x_k - x_{k-1} \]  \hspace{1cm} (4.26)

which simplifies the coder operation and reduces the average number of instructions. As can be expected, \( d_k \) is characterized by a skew probability distribution, which enables effective entropy coding. This probability distribution is shown in Fig. 4.13 for the sequence BOARD. The entropy \( H(d) \), calculated according to:
4.3 An effective algorithm for low frame-rate video coding

\[ H(d) = \sum_{i=1}^{16} p(d_i) \log_2(1/p(d_i)) \]  

is equal to 1.37 bit/pel. This figure is a lower bound for \( R \) and can be approximated by entropy coding of the DPCM output. Note that entropy values will be used for comparison throughout this section. As can be seen in Fig. 4.13, the probability that \( d \) lies in the range [-1, ..., +1] appears to be 98% (BOARD). Hence, application of a 3- or a 2-bit quantizer will only slightly reduce the coding rate \( R \), but will degrade edge response somewhat. One-bit quantizers (delta modulation) will not be considered because of clearly visible artefacts like slope overload, and granular noise. Thus it is defendable to limit ourselves to 4-bit quantizers. Expression (4.26) requires 3 ins/pel. Decoding is possible with the same amount of instructions. It is supposed that the entropy coding step is implemented according to the Modified Tunstall Coding (MTC) algorithm. In Algra [92a] and Chapter 5 it is shown that MTC is a fast and efficient entropy coding method, which outperforms Huffman coding with respect to complexity and coding rate in the case of a skew source with a limited alphabet. MTC encoding and decoding can be carried out with 6.5 and 5.75 ins/pel, respectively. In other words, the complete DPCM coding operation can be performed with 9.2 ins/pixel.

b) DPCM 3d and 4th-order, 2-dimensional DPCM

Two higher-order DPCM schemes were evaluated which both exploit 2-dimensional redundancies. The coding algorithms:

\[ d_k = x_k - (0.5A + 0.125B + 0.25C + 0.125D) \]  

and

\[ d_k = x_k - (0.5A + 0.25C + 0.25D) \]

The results of encoding HEADS and BOARD are listed in Table 4.4. A gain in compression ratio is achieved at the cost of an increase of computational complexity. The fastest method is to use a look-up table for prediction calculation. Assuming again 4-bit linear quantization, the 4th order encoder algorithm is realisable by 10 ins/pel (3d order: 8). However, the prediction calculation requires a 64 kbyte or a 16 kbyte table in memory for the 4th and 3rd order cases, respectively. The 4th order decoder applies the same prediction table and requires also 10 ins/pel (3d order: 8).
c) ADPCM
One DPCM scheme with adaptive prediction was evaluated. The algorithm of (4.30) resembles third order two-dimensional adaptive prediction:

\[ d_k = x_k - \text{Med} \{ A, C, 0.5A + 0.25C + 0.25D \} \] (4.30)

where \( \text{Med} \{ p, q, r \} \) denotes the (running) median value of \( p, q \) and \( r \). A similar look-up table technique as used with b) leads to 8 ins/pel and requires a 4096 byte table in memory. The advantage of median prediction is that the decoder does not need additional predictor status information, as it performs the same prediction calculation as the encoder and so requires the same computational effort.

4.3.3.3 Vector Quantization

A demanding criterion for application of VQ is that the encoder search process should be made sufficiently computation efficient. A binary-tree structured codebook (BTBQ) reduces the search time significantly, but not enough with respect to our criteria: \( 2\log_2 c \) multiplications and operations per pixel are required. Translated into the number of required assembler instructions per pixel:

\[ I > 6 \log_2 c \text{ ins/pel} \] (4.31)

are needed to determine the correct output vector. For a typical codebook size of \( c = 256 \), this leads to 48 ins/pixel. Even if a very small codebook would be used, e.g., in combination with the product code method, then the number of instructions/pixel is unacceptably high (for example: \( c = 16 \) results in 24 ins/pixel).

An alternative approach is the implementation of a full search codebook by a look-up table. Let the vector elements address an \( N \)-dimensional array, containing all \( c \) output vector codes, then a few simple operations are necessary to determine the correct codeword. However, the memory occupation \( M \) of such an array limits the dimension \( N \). If the vector elements are represented in \( n \) bits, then

\[ M = 2^n \text{ bytes} \] (4.32)

Memory occupation of the codec should be < 512 kbyte. Where \( N \geq 4 \), it follows that \( n = 4 \) and \( N = 4 \), implying VQ carried out on 2x2 pixel blocks and a memory occupation of 64 kbyte. A way to exploit the correlation between neighbouring 2x2 blocks is the application of 2x2 MSVQ followed by a DPCM operation on the means of the blocks:
4.3 An effective algorithm for low frame-rate video coding

Thus each vector $x$ results in two values: a DPCM-coded mean ($y_1$) and a vector-quantized residual ($y_2$). The results of a complexity analysis (Table 4.2) show that, on average, only 7.1 ins/pixel are required, including entropy coding. This method will be referred to as FEVQ (for Fast Entropy coded Vector Quantization) further on, for convenience.

The quality of a VQ coder depends on the quality and structure of the codebook. The usual approach to design a VQ coder is iterative improvement of the codebook by processing a set of training images, which are representative for the application. However, in our case the size of the codebook is small. If reasonable compression is to be achieved before the entropy coding step, and given that the representation of the block mean $\mu$ requires 4 bits, the VQ output vector should be represented by much less than 12 bits. Hence, the codebook size $c$ is much less than 4096. $c$ should be as small as possible, with acceptable image quality. Small codebooks, e.g., $c = 32$, enable a non-iterative approach to obtain the best coding vector set. Chen & Bovik [90] proposed a similar method called visual pattern image
Table 4.2 Complexity analysis of FEVQ codec in average number of assembler instructions per block (2x2 MSVQ/DPCM)

<table>
<thead>
<tr>
<th>encoder:</th>
<th>I (ins/block)</th>
<th>decoder:</th>
<th>I (ins/block)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean calculation:</td>
<td>6</td>
<td>MTC decoding:</td>
<td>5.75</td>
</tr>
<tr>
<td>DPCM on mean:</td>
<td>3</td>
<td>DPCM decoding:</td>
<td>3</td>
</tr>
<tr>
<td>MTC on DPCM value:</td>
<td>6.5</td>
<td>MTC decoding:</td>
<td>5.75</td>
</tr>
<tr>
<td>Residual VQ:</td>
<td>7</td>
<td>Block composition:</td>
<td>13</td>
</tr>
<tr>
<td>MTC on VQ value:</td>
<td>6.5</td>
<td>Total:</td>
<td>27.5</td>
</tr>
<tr>
<td>Total:</td>
<td>29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

coding (VPIC), which exploits the properties of the HVS in an original way. In the basic VPIC scheme each 4x4 block is classified as a uniform block (no edge apparent) or as one of 14 edge-type blocks. The latter only consist of edges with horizontal, vertical or diagonal orientations. In the case of a uniform block only the mean of the 16 pixels is coded. In the case of an edge-type block, also the edge type and the edge gradient (i.e., the luminance gradient rectangular oriented with respect to the edge) are coded. Finally, 1 bit/block is allocated to indicate whether the block is uniform or not. Compression results are comparable with other VQ methods based on 4x4 blocks \( = 0.5 - 1.0 \text{ bit/pel} \). But with respect to computational complexity VPIC outperforms all other VQ methods: the authors report 2.25 additions, 0.53 comparisons and 0.19 multiplications per pixel. Our estimation of these figures in terms of 286 assembler instructions results in 9 ins/pel.

Although VPIC is unique in the research literature with respect to complexity, the following reasons force us to develop an alternative method based on FEVQ:

1) The limitation to 14 edge patterns for a 16-pixel area significantly reduces the image resolution with respect to the original. Consequently, edge lengths must be multiples of 4 pixels (block width). Thus, a line width of 2 pixels cannot be displayed, and objects with a dimension of 3 pixels or less become blurred.

This resolution degradation is confirmed by examination of the pictures of reconstructed images presented in Chen and Bovik's paper. Considering tele-education applications, e.g. images of a white-board, where often narrow lines are displayed, this property is undesirable.

2) FEVQ actually splits up the image into a LF component, resembling 50% of the resolution of the original, and a HF part. This construction opens a perspective for a two stage progressive image build-up feature, with a significant further reduction of
the image transfer time for the medium resolution image.

3) FEVQ achieves comparable or even slightly better computational efficiency.

Still, the approach of the decomposition of an image into a set of visual patterns, resembling (sharp) edge orientations and uniform blocks, will be adopted from Chen and Bovik [90]. The advantage of their method is that, firstly, such a coding scheme is independent of a certain training set, and secondly, it is better adapted to the properties of the HVS. Our method decomposes 2x2 blocks into uniform and edge-type blocks. Edge-type blocks are classified into 4 edge patterns, depicted in Fig. 4.15. With each edge-type block, also the edge gradient is coded. Note that, when mapped to a 4x4 pixel area, this decomposition allows at least $5^4 = 625$ different patterns, guaranteeing better resolution than VPIC. If the edge gradient is coded into 3 bits (resembling 8 levels) similar to Chen & Bovik [90], the total amount of different residual 2x2 images is $1 + 2^5 = 33$. In other words, the VQ codebook size $c$ equals 33.

Table 4.3 lists the set of visual edge type patterns which were used for experimental verification (algorithm 1). Note that the set size is reduced to 31, resulting in a more convenient codebook size of 32. Jointly coding of gradient index and edge orientation is easily obtained by a look-up table, with the advantages of a reduced number of instructions per pixel, and the possibility to add patterns which cannot be constructed with independent coding of gradient and orientation. The first row of Table 4.3 (gradient index 0) is a typical example of such a pattern, added in order to obtain a precise reconstruction in areas with minor luminance variation. However, experiments were carried out with a somewhat reduced set: here the first row was deleted and thus $c = 28$ (algorithm 2). The function $Q(x)$ is implemented by a 4-dimensional array, used as a look-up table. The array contains output vectors $y$ for each possible input vector $x$, calculated according to (4.22), where
Table 4.3 Visual edge type patterns (FEVQ algorithm 1)

<table>
<thead>
<tr>
<th>Gradient Index</th>
<th>Orientations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>+1</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>+1</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>+1</td>
</tr>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>+2</td>
</tr>
<tr>
<td></td>
<td>+2</td>
</tr>
<tr>
<td>4</td>
<td>+3</td>
</tr>
<tr>
<td></td>
<td>+3</td>
</tr>
<tr>
<td>5</td>
<td>+4</td>
</tr>
<tr>
<td></td>
<td>+4</td>
</tr>
<tr>
<td>6</td>
<td>+5</td>
</tr>
<tr>
<td></td>
<td>+5</td>
</tr>
<tr>
<td>7</td>
<td>+7</td>
</tr>
<tr>
<td></td>
<td>+7</td>
</tr>
</tbody>
</table>

\[ d(x,y) = \sum_{i=1}^{c} |x_i - y_i|^2 \]  

Table 4.4 gives the average results of \( R \) in bit/pel. Compare also the Figures 4.16 - 18, where a selection of the original and reconstructed images of the sequences BOARD and HEADS, and of the image LENA is shown. Algorithm 2 results in somewhat higher compression, but is affected by more pronounced contouring effects.

Comparing the results of the various algorithms (Table 4.4), it is clear that FEVQ results in the fastest end-to-end transfer times. Note, however, that practical values for \( R \) will be slightly higher than those listed in Table 4.4, since entropy coders do not realize a coding efficiency of 100%. A considerable additional reduction of \( T_f \) can be realized if FEVQ is combined with CR.
Fig. 4.16  

a) Original frame of BOARD (512 x 384 pixels) displayed at 4 bit/pixel  
b) Reconstructed frame, with FEVQ-1 encoding algorithm (coding rate: 0.76 bit/pixel)  
c) Reconstructed frame, with FEVQ-2 encoding algorithm (coding rate: 0.47 bit/pixel)
Fig. 4.17  a) Original frame of HEADS (128 x 96 pixels) displayed at 4 bit/pixel  
b) Reconstructed frames, with FEVQ-1 encoding algorithm (coding rate: 1.00 bit/pixel)  
c) Reconstructed frames, with FEVQ-2 encoding algorithm (coding rate: 0.70 bit/pixel)
Fig. 4.18  a) Original frame of LENA (512 x 384 pixels) displayed at 4 bit/pixel  
b) Reconstructed frame, with FEVQ-1 encoding algorithm (coding rate: 1.06 bit/pixel)  
c) Reconstructed frame, with FEVQ-2 encoding algorithm (coding rate: 0.81 bit/pixel)
Table 4.4  Coding rates (bit/pel) and complexity (ins/pel) for various methods. $T_f$ yields for $L = 512x384$, $c_i = 0.15$, $B = 9600$ bit/s

<table>
<thead>
<tr>
<th>Method</th>
<th>BOARD</th>
<th>HEADS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R$</td>
<td>MSE</td>
</tr>
<tr>
<td>DPCM 1-ord</td>
<td>1.37</td>
<td>4.62</td>
</tr>
<tr>
<td>DPCM 3-ord</td>
<td>1.21</td>
<td>4.62</td>
</tr>
<tr>
<td>DPCM 4-ord</td>
<td>1.22</td>
<td>4.62</td>
</tr>
<tr>
<td>ADPCM</td>
<td>1.25</td>
<td>4.62</td>
</tr>
<tr>
<td>FEVQ-1</td>
<td>0.76</td>
<td>6.24</td>
</tr>
<tr>
<td>FEVQ-2</td>
<td>0.47</td>
<td>7.36</td>
</tr>
</tbody>
</table>

4.3.4 Conditional Replenishment (CR)

With conditional replenishment, most transmission capacity is used only to replenish those areas of the picture which have changed significantly since the previous frame. Differences between successive frames are caused by moving objects, camera noise and the camera’s automatic gain control (AGC). An effective CR algorithm should discriminate between motion and non-motion caused differences. Noise can be suppressed to a large extent by temporal and/or spatial noise filtering. The identification of AGC effects is only possible by complex image processing and will not be covered here. It has been observed that in pictures contaminated by white Gaussian noise where the signal-to-noise ratio is greater than 40 dB, frame-to-frame differences larger than 3 percent of maximal signal amplitude are almost certainly due to movement (Haskell et al. [72]). When 4-bit signal quantization is used prior to the CR operation, the minimal detection threshold resembles $2^{-4}$ of the maximal amplitude (6.7 percent). Hence noise-induced frame-to-frame differences $> 2^{-4}$ are negligible. This is experimentally verified by measuring the frame-to-frame differences of the sequence BOARD-STL, which consisted of 4 subsequent frames, but without motion (see Table 4.5). However, differences equal to $2^{-4}$ of the maximal amplitude (corresponding with 1 L.S.B. in 4-bit uniform quantization) may be caused by noise as well as motion. Hence, an effective motion detector should discriminate between noise-induced differences and real movement. One way is to suppress the noise by spatial filtering prior to the CR operation. A fairly simple method to remove changes due to noise is to ignore any changes immediately preceded and followed by two unchanged elements (Haskell et al. [72]). In Limb & Pease [71] a spatial filter is
Table 4.5 Probability distributions of frame-to-frame differences of a sequence with motion (BOARD) and without motion (BOARD-STL) based on 2x2 blocks

<table>
<thead>
<tr>
<th>Condition</th>
<th>Category</th>
<th>BOARD-STL</th>
<th>BOARD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta_k = 0$</td>
<td>2</td>
<td>0.336</td>
<td>0.235</td>
</tr>
<tr>
<td>$\Delta_k = 0$</td>
<td>3</td>
<td>0.147</td>
<td>0.065</td>
</tr>
<tr>
<td>$\Delta_k = 1$</td>
<td>3</td>
<td>0.405</td>
<td>0.290</td>
</tr>
<tr>
<td>$\Delta_k = 2$</td>
<td>3</td>
<td>0.098</td>
<td>0.135</td>
</tr>
<tr>
<td>$\Delta_k = 3$</td>
<td>3</td>
<td>0.011</td>
<td>0.062</td>
</tr>
<tr>
<td>$\Delta_k = 4$</td>
<td>3</td>
<td>0.001</td>
<td>0.036</td>
</tr>
<tr>
<td>$\Delta_k \geq 2$</td>
<td>1</td>
<td>0.003</td>
<td>0.176</td>
</tr>
</tbody>
</table>

applied, giving the average over four adjacent elements along a line.

Let $d_{ik}$ denote the difference of pixel $i$ of block $k$ of the current frame and the corresponding pixel in the previous frame. Then $\delta_k$ and $\Delta_k$ are defined for a $N$ pixel block:

$$\begin{align*}
\Delta_k &= \text{MAX} \{ |d_{ik}|, \ i = 0, \ldots, N-1 \} \\
\delta_k &= \sum_{i=0}^{N-1} d_{ik}
\end{align*}$$

(4.35)

In Table 4.5 blocks are compared and classified into three categories: 1) $\Delta_k \geq 2$ (represents motion); 2) $\Delta_k = 0$ (represents no motion), 3) $\Delta_k = 1$ (motion- or noise-induced change).

If category 3 is further split up into subcategories representing various values of $|\delta_k|$, a method is provided for a fast and efficient estimation of whether the category 3 change is due to motion or to noise: The block will be classified as a changed block when the joint condition ($\Delta_k \geq 2$) $\lor$ ($|\delta_k| \geq 3$) is fulfilled. Note that $\delta_k$ resembles the mean of the block, which implies that the subsequent mean calculation by the intra-frame algorithm can be skipped. Analysis of the category 1 blocks revealed that 99% of these blocks also fulfilled the condition $|\delta_k| > 2$. Hence just checking the latter condition is sufficient. This results into the coding algorithm given in Table 4.6. $Q[.]$ represents the 64 kbyte look-up table, as used for the VQ encoding step. $M[k]$ is a $L/4$ byte frame buffer. The generated data consist of DPCM/mean ($y_j$) codewords, VQ codewords ($y_2$), and a bitmap (Cmap).

Fig. 4.19 depicts the fraction of frame-to-frame differences $m$, measured according to this criterion for BOARD. Table 4.7 lists the resulting average coding rates of FEVQ-I and FEVQ-2 in combination with CR. These values include a straightforward block addressing scheme consisting of a separately transmitted bitmap (size: $L/N$ bits) where each bit indicates
whether a block was replenished or not. Let \( m_N \) denote the fraction of changed \( N \)-pixel blocks, then the resulting coding rate is expressed by

\[
R'_N = \frac{1}{N} + m_N R
\]

(4.36)

Fig. 4.19 Measured fractions of frame-to-frame differences for the image sequence "BOARD"

Table 4.6 FEVQ algorithm with conditional replenishment for a frame of \( L \) pixels

```
FOR k=0 TO (L/4 - 1) DO
    \( \mu_k := (x_{0k} + x_{1k} + x_{2k} + x_{3k} + 1) \text{ DIV } 4 \)
    IF \( \mu_k \neq M[k] \) THEN
        \( y_{1k} = (\mu_k - M[k-1]) \text{ MOD } 16 \)
        \( y_{2k} = Q[x_{0k},x_{1k},x_{2k},x_{3k}] \)
        \( \text{CRmap}[k] = 1 \)
        \( M[k] = \mu_k \)
    ELSE \( \text{CRmap}[k] = 0 \)
```
It is worthwhile investigating whether 2x2 block clustering effects could be exploited for further reduction of the coding rate. Assume therefore that the CR decision would be based on \( N \) pixel blocks, with for convenience \( N = 2^{2n}, n = 1,2,... \), and where such an \( N \)-block is replenished only if one or more 2x2 blocks are categorized as changed. Applying this criterion, the fractions \( m_N \) of changed \( N \)-blocks of the sequence \( BOARD \) were measured for \( N=16 \) and \( N=64 \). Fig. 4.20 depicts the resulting \( R' \) as a function of \( R \). Obviously some further reduction of the coding rate is possible.

**Table 4.7 Coding rates (bit/pel) for FEVQ, with and without conditional replenishment**

<table>
<thead>
<tr>
<th>Method</th>
<th>( R ) (bit/pel) BOARD</th>
<th>( R ) (bit/pel) HEADS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEVQ-1</td>
<td>0.76</td>
<td>1.00</td>
</tr>
<tr>
<td>FEVQ-2</td>
<td>0.47</td>
<td>0.70</td>
</tr>
<tr>
<td>FEVQ-1 + CR</td>
<td>0.49</td>
<td>0.61</td>
</tr>
<tr>
<td>FEVQ-2 + CR</td>
<td>0.40</td>
<td>0.54</td>
</tr>
</tbody>
</table>

An alternative approach to improve the compression effectiveness is the application of a second level of conditional replenishment. Assume that the image of \( L \) pixels is subdivided into \( L/L_p \) images of \( L_p \) pixels. Then second-level CR involves that a sub-image is not transmitted at all if it contains less than a certain number of changed 2x2 blocks. For this purpose the data of a coded and transmitted sub-image is accompanied by a header, containing a sub-image identifier. When the decision whether to replenish a sub-image is based on a sample check, the speed of the coding process is further improved. For the decision criterion and the value of \( L_p \) refer to section 4.4, where coding rate values of such a system are also reported.
Fig. 4.20 Effective coding rate $R'$ (bit/pel) as a function of the intraframe coding rate $R$ for conditional replenishment based on blocks of $N$ pixels
Fig. 4.21  a) Original frame of BOARD (512 x 384 pixels) displayed at 4 bit/pixel  
b) Reconstructed frame, with FEVQ-1 and CR encoding algorithm (coding rate: 0.49 bit/pixel)  
c) Reconstructed frame, with FEVQ-2 and CR encoding algorithm (coding rate: 0.40 bit/pixel)
4.4 LFVC codec design and performance

In order to verify the results of Section 4.3, a complete software-based codec has been realized, based on FEVQ, with options for single and two-level conditional replenishment. Where critical, system routines were optimized with respect to speed. Fig. 4.22 reflects the overall software architecture. The system is able to process image sequences, acquired from a frame grabber or retrieved from a file. The decoded and reconstructed images can be displayed on a VGA screen in 16 grey levels. For evaluation purposes, the system calculates and displays the current coding rate, and the encoding and decoding times.

![System architecture diagram](image)

*Fig. 4.22 System architecture*

4.4.1 Design considerations and results

4.4.1.1 Intermediate and encoded data formats

According to the conditional replenishment algorithm as given in Table 4.6, both encoding as well as decoding requires a frame buffer containing the block means of the latest frame. Hence the buffer size should be $L/4$ bytes (assuming 1 byte/mean). However, in order to enable immediate exposure of the most recent LFVC image, the high resolution content of
the images should also be available in the frame memory. This requirement follows from
typical tele-education situations, where the teacher should have the possibility to switch
between the following screen configurations: 1) presentation of text, graphics etc. in the
shared workspace and simultaneously display of a small format LFVC image inserted at a
corner of the screen to observe the class room situation; 2) display of the large format LFVC
image, for a closer watch into the class, with consequently the shared workspace obscured.
Hence, in order to avoid unacceptable response times, large format LFVC (de)coding should
be executed continuously, even during the small (or medium) format mode. In other words,
the buffer should include the residual vector quantization data of the image (5 bits per block).
Reserving 4 bits per block for future colour encoding, then 13 bit/block need to be stored.
The resulting format is depicted in Fig. 4.23-a. To simplify storage and readout, 2 bytes per
block are used.

The encoded data format is based on representing sub-images independently of each
other, in order to facilitate second-level CR. Fig. 4.23-b shows the encoded data format of
one sub-image of \( L_p \) pixels. For \( L_p \) the value 12288 is chosen; this is equal to the number
of pixels of a small format image (128x96). This simplifies the implementation of some data
management functions related to format switching. The header contains a sub-image
synchronization code followed by the identifier (0..15), and 2 pointers. The header is
followed by the fixed-size CR bit map (384 byte) and 2 variable-size fields: encoded
mean/DPCM and encoded residual VQ information. The first pointer addresses the encoded
residual VQ field, and the second pointer gives the first position of the next sub-image.
4.4.1.2 Intraframe encoding and CR

The FEVQ-1 algorithm was implemented, albeit with the following modification. Experiments with the sequence BOARD revealed that visual edge-type patterns with gradients 5,6 and 7 did not occur at all (refer to Table 4.3). Hence, the codebook was further reduced to 21 vectors. The MTC segment sets included were based on probability distributions of $y_1$ and $y_2$, as obtained from 6 different images (including the average values of BOARD).

Two versions of the system were realized: 1) FEVQ with conditional replenishment according to Table 4.6 (FEVQ-CR); 2) FEVQ-CR with additional second-level CR, denoted by FEVQ-CR$.^2$. The implemented FEVQ-CR$.^2$ version involves a sample check (based on a 1:4 subsampling grid) on 192 block means, prior to the encoding of the sub-image. Encoding and transmission of the sub-image is performed only if one or more of the mean values differ more than 1 from their temporal predecessors.

<table>
<thead>
<tr>
<th>Table 4.8 Results of Low Frame-Rate Video Coding of sequence BOARD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R$ (bit/pel)</td>
</tr>
<tr>
<td>$I$ (ins/pel)</td>
</tr>
<tr>
<td>$T_{TF}$ (s)</td>
</tr>
<tr>
<td>$R = 9.6$ kbit/s; $c_I = 0.15$</td>
</tr>
<tr>
<td>$T_{TF}$ (s)</td>
</tr>
<tr>
<td>$R = 56$ kbit/s; $c_I = 0.15$</td>
</tr>
<tr>
<td>$T_{TF}$ (s)</td>
</tr>
<tr>
<td>$R = 56$ kbit/s; $c_I = 0.3$</td>
</tr>
</tbody>
</table>

4.4.1.3 Results

The average coding rate for BOARD appears to be 0.68 and 0.38 bit/pel for FEVQ-CR and FEVQ-CR$.^2$, respectively. These values are slightly higher than those presented in Section 4.4, due to the entropy coding process, which does not achieve a coding efficiency of 100 percent. Table 4.8 gives the average transfer times for BOARD as obtained for 9.6 and 56 kbit/s (with a CPU-load of 30 and 60%, corresponding to $c_I = 0.15$ and 0.3 respectively). On the average, $T_{TF}$ is less than 10 seconds, which is acceptable. For medium and small images these values are reduced to (estimated) below 2.5 and 0.6 seconds, respectively.

Note that at a CPU-load of 30%, the transfer time is not influenced by the bitrate (ranging from 9600 to 56 kbit/s), because the processing time of the terminal is dominant.
Fig. 4.24 End-to-end transfer times of sequence BOARD (512x384 pixels), excl. frame-
grab and display times, for FEVQ-CR and FEVQ-CR²

4.4.2 Channel errors

Highly compressed image data is inherently sensitive to bit errors arising in transmission. Especially prediction and entropy coding methods suffer. Obviously, a bit error in the LFVC sub-image header information will be disastrous and probably lead to loss of sub-image frame synchronization. An erroneous CR bit may cause severe corruption of the image: not only the 2x2 block corresponding to the false bit is wrongly displayed, but all subsequent information on changed blocks within the current sub-image will undergo a position shift. Although a bit error in received MTC data will never result in loss of word synchronization (unlike Huffman encoded data, see Chapter 5), two effects may appear: a) One or more source-letters become corrupted; b) The decoder may misinterpret the number of source-
letters, resulting in a position shift of the succeeding blocks of the sub-image. Furthermore the replenishment mechanism implies that the effect of an error may remain visible for a rather long time.
In general, three methods have been considered to reduce the effects of transmission errors: 1) Restriction of the damage caused by errors to as small an area as possible, 2) Detection of errors and retransmission of corrupted blocks using an Automatic Retransmission Request (ARQ) System and 3) Detection of errors and their correction at the receiver using a Forward Error Correcting (FEC) code (Netravalli & Haskell [87]).

In the case of LFVC a combination of these methods would be preferable. Restriction of damage to a small area (spatially and temporally) can be achieved by the intersection of the data through specific framing bytes, marking certain image area borders, and by a periodic forced refreshing mechanism (Candy et al. [71]), resulting in the replenishment of the complete image during a time interval of several frames. Depending on the bit error rate (BER) of the channel, a trade off exists between the increased coding rate and the resulting image quality. For example, sub-division into areas of 128 x 2 pixels requires 24 framing bytes per sub-image, corresponding to an average increase of the coding rate of about 3%. A refreshment-period of 20 frames implies an additional increase of 10%. Then, the effect of a bit error is limited to (average) 0.065 % of the image area.

FEC methods are widely applicable, also in situations where ARQ is not appropriate, such as multicasting and long-delay (satellite) links. For example, the (120,128) extended Hamming code is a FEC code, which is suitable for fast software implementation (refer to Appendix A). This code is able to correct 1 and to detect 2 bit errors in a 120 bits block. When the FEC decoder detects 2 errors in a block, a simple error concealment approach can be employed, by not updating the corresponding image area at all. Hence, disastrous image damage only occurs if at least 3 bit errors appear in a block. In this way acceptable image quality can be maintained for BER rates up to $10^{-4}$, unless occasional burst errors occur.

4.4.3 Concluding remarks

It has been shown that the developed FEVQ-CR and FEVQ-CR² algorithms are suitable for software implementation and result in acceptable end-to-end transfer times, even for large format images.

An additional reduction of the effective transfer time is possible, when the system is implemented according to a two-stage progressive build-up scheme. The encoding process can easily be re-ordered in the sense that as a first step each sub-image is encoded and transmitted based on 50% resolution. This involves conditional replenishment processing and block mean encoding. Secondly, the high resolution part of the image (the residual VQ encoding) is encoded and transmitted. From the complexity analysis as given in Table 4.4, it can be estimated that the medium resolution part of the image is available at the receiver.
4.4 LFVC codec design and performance

within 5 seconds (large format, CPU-load = 30%). Note that the qualification 'medium resolution' corresponds to 320 x 240 pixels, which is comparable to Digital Video Interactive's (DVI) full screen motion video resolution of 256x240 (Fox [89]). As can be observed from Fig. 4.23, the adopted encoded data format lends itself to an easy split up into two separate frames, where the first one contains CR and mean information and the second contains the residual VQ data.

The proposed system is appropriate for extension to colour (de)coding. Assuming 16 available basic colours (in combination with 16 luminance levels, yielding 256 'colours'), and allowing 2:1 subsampling for the colour information (This does not affect the subjective image quality, see Netravali & Haskell [88], p.437); one 4 bit value per 2x2 block is sufficient. Further compression of these data can be achieved by application of a coding scheme similar to the luminance mean values: DPCM followed by MTC. Preferably a VGA compatible frame-grabber should be selected, which simplifies colour identification. The encoded data format is then extended with a colour field and an extra pointer in the header. Determination of the 16 basic colours palette is a computationally intensive operation (e.g., see Robinson et al. [91]). However the palette does not need frequent updating, due to the static nature of the typical tele-education scene.

Considering the continuously increasing clock speed of subsequent CPU generations, one might query whether the LFVC method will still remain effective compared to competing software-based schemes. The following considerations show that LFVC performs better than

\[ B = 9600 \text{ bit/s} \]

\[ \text{CPU clock (MHz)} \]

\[ \text{Transfer time (s)} \]
Fig. 4.26 Comparison of transfer times as a function of CPU clock rate for LFVC and JPEG DCT, $B = 56$ kbit/s, CPU-load = 30%

e.g., the fast JPEG method (Arai et al. [88]), even if fast PCs are used. The Figures 4.25 and 4.26 compare the transfer times (from eqn. 4.11) of the JPEG method with LFVC as a function of the CPU clock speed ($f_{cpu}$), assuming $R = 0.5$ for both compression schemes and $i_r = 12/f_{cpu}$. Obviously LFVC is the preferable method for processor speeds up to 40 MHz, in the case of 9600 bit/s. At higher bitrates LFVC outperforms the JPEG method even if the fastest CPU’s are applied. However, this advantage is reduced when processors with a cache memory (such as the 80486) are applied, which show instruction durations, on the average, much less than $12/f_{cpu}$.

4.5 Summary and conclusions

In this chapter a Low Frame Rate Video Coding method (LFVC) was presented, which has been developed as a sub-system for a real-time tele-education terminal. It encompasses a fast and efficient coding scheme, appropriate for software implementation, and aimed at narrow bandwidths. The coding scheme is useful for use in a multimedia tele-education terminal, where a requirement exists for implementation on standard microcomputer equipment without dedicated hardware. Such a subsystem allows the teacher to watch the classroom (and vice versa) during the session, by near full-screen image sequences.
4.5 Summary and conclusions

An analysis on the relations between coder complexity, compression factor, CPU-load and transfer time has been performed, revealing the need to develop a coding scheme with significant lower complexity than present standard methods such as JPEG DCT. A survey of intraframe compression methods resulted in the initial selection of a number of DPCM methods and a VQ scheme for this purpose.

From experiments and complexity analyses, the FEVQ (Fast and Efficient Vector Quantization) algorithm resulted, applying LUT-based MSVQ on 2x2 pixel blocks in combination with DPCM on the block means and Modified Tunstall Coding. Low-complexity interframe compression can be achieved by conditional replenishment. A two-level approach was adopted: replenishment of 2x2 pixel blocks and of 128x96 pixel blocks (called: sub-images). With sub-image replenishment based on a sample check mechanism, a very efficient encoder implementation is possible.

In order to verify the results obtained, a complete software-based codec has been realized, achieving acceptable image end-to-end transfer times even at low data rates of the transmission channel (9600 bit/s). The encoding process is appropriate for extension to colour encoding. Furthermore it can easily be re-ordered into a two-stage progressive transmission scheme to reduce the effective transfer time by 50%.
References in Chapter 4


Algra, T. [92b], "Low frame rate video coding for small bandwidth realtime multimedia teleeducation systems", *Proc. IEEE Africon'92*, pp.352-357, Sept. 1992


Candy, J.C., M.A.Franke, B.G.Haskell, F.W.Mounts [71], "Transmitting television as clusters of frame to frame differences", *Bell System Techn. Journal*, pp.1889-1918, 1971


Haskell, B.G., D.J.Connor, F.W.Mounts [73], "A frame-to-frame picture phone coder for signals


Koga, T. et al. [81], "Motion compensated interframe coding for videoconferencing", *Proc. NTC 31, New Orleans*, Dec 1981


Liou, M.L. [91], "Overview of the px64 kbit/s video coding standard", *Comm. of the ACM*, Vol.34, No.4, pp.59-73, 1991


No.6, pp.637-645, 1987

Mussmann, et al. [85], "Advances in picture coding", Proc. IEEE, Vol.73, No.4, 1985


Vetterli, M., H.J.Nussbaumer [84], "Simple FFT and DCT algorithms with reduced number of operations", Signal Processing, 6, pp.267-278, 1984


Ziv, J. [90], "Variable-to-fixed-length codes are better than fixed-to-variable-length codes for Markov sources", IEEE Trans. Inform. Theory, IT-36, No.4, pp.861-863, 1990
5 FAST AND EFFICIENT VARIABLE-TO-FIXED-LENGTH CODING

5.1 Introduction

More than a quarter of a century ago (1952) Huffman [52] introduced a minimum redundancy method of source encoding now known as a 'Huffman code'. Recent work on fast and efficient Huffman encoding and decoding algorithms can be found in Tanaka [87], Sieminsky [88], Hirschberg & Lelewer [90]. Tanaka [87] introduced high-speed encoding and decoding using a finite-state automaton method. Sieminsky [88] applied a similar approach to fast Huffman decoding with directions for decoder complexity analyses. The drawback of this finite-state automaton approach is the large amount of memory required for the decoding process. Hirschberg & Lelewer [90] proposed a number of alternatives with reduced memory requirements, however at the cost of speed.

The Huffman code belongs to the general class of Fixed-to-Variable-length coding schemes (FV), which assign to every source-letter a codeword of a different length. For this class the optimal code is generated by the Huffman procedure (Huffman [52]), i.e., for a given source it minimizes the average codeword length. Thus encoder and decoder implementations have to do bit manipulations, a property resulting in complex algorithms especially in the case of software decoding (Sieminsky [88], Hirschberg & Lelewer [90]).

For some frequently used classes of discrete memoryless sources, Variable-to-Fixed-length coding (VF) methods appear to be more attractive. VF codes for discrete memoryless sources were investigated by Tunstall [67], Jelinek and Schneider [72]. Tunstall [67] found a procedure for constructing the optimal segment set for a discrete memoryless source, in the sense that it minimizes the coding rate for a given number of segments. Jelinek & Schneider [72] proved that the code rate approaches the source entropy if the number of segments in the set grows to infinity. Tjalkens & Willems [87] investigated variable-to-fixed-length
coding algorithms for Markov sources. Liu [92] proposed a variable-to-fixed-length coding algorithm for application with Differential Chain Coding of line drawings, because of the resistance against word synchronisation loss due to transmission errors, compared to Huffman codes. Ziv [90] proved theoretically that for kth-order Markov sources, a variable-to-fixed-length code performs better than the best fixed-to-variable code (Huffman code).

This chapter shows that for discrete memoryless sources, characterized by a skew probability distribution (and consequently a low source entropy value), it is advantageous to apply VF codes. VF codes perform better than FV codes with respect to the coding rate (depending on the source statistics), complexity, and sensitivity to transmission errors. Furthermore a modified version of Tunstall’s algorithm is proposed, with a coding rate which converges faster to the source entropy.

Section 5.2 introduces Tunstall’s algorithm, and analyses the advantages of Tunstall coding over Huffman coding with respect to coding rate, transmission error sensitivity, and complexity. Section 5.3 introduces Modified Tunstall Coding, and deals with coding rate performance and complexity aspects of this new type of VF coding. Furthermore the sensitivity to variations of source statistics is analysed.

Some key parts of this chapter have been published previously in Algra [92].

5.2 Tunstall codes

5.2.1 Introduction to Tunstall codes

We consider a discrete memoryless source (DMS) with source letter alphabet \( X_c \) and probability distribution \( Q(x) \), \( x \in X_c \).

\[
X_c = \{0,1,2,..c-1\}
\]

The convention will be adopted that

\[
Q(0) \geq Q(1) \geq .. \geq Q(c-1)
\]

(5.1)

The entropy of the source is

\[
H(Q) = \sum_{x \in X_c} -Q(x)\log_2 Q(x)
\]

(5.2)
In a variable-to-fixed (VF) length coding situation the encoder chops the source output sequence into segments of about the same probability. To inform the decoder, the encoder sends it the lexicographical index of the segments. The decoder is then able to reconstruct the segments. The segments belong to a segment set $S$ with cardinality $T$.

$$s_0, s_1, \ldots, s_{T-1} \in S$$

Let $L(s)$ be the length of segment $s$ (number of source-letters). The $s$ transmitted by the $T$-ary source are independent and identically distributed random variables. The distribution $Q(x)$ generates the following distribution on the alphabet $S(T)$. If $s = x_1 \ldots x_{L(s)}$ then the probability of $s$ is:

$$P(s) = \prod_{i=1}^{L(s)} Q(x_i)$$

(5.3)

Let

$$\bar{L} = \sum_{s \in S} P(s)L(s)$$

(5.4)

Consider $s_i \in S(T)$. Let us assign a binary sequence $\Theta(s_i)$ to $s_i$ with the following property. $\Theta(s_i)$ has length $n = \lceil \log_2 T \rceil$ (the symbol $\ceil{y}$ stands for 'smallest integral value greater than or equal to $y$').

A measure of the effectiveness of a VF coding scheme is the coding rate:

$$R = n/\bar{L} \text{ bit/source-letter}$$

(5.5)

Tunstall [67] proved that the following algorithm obtains for every admissible size $T = c + (k-1)(c-1)$ the complete and proper segment set $S(T)$ minimizing the coding rate.

**Algorithm 1:**

Let $S(c) = X_c$. This is the complete and proper segment set of size $c$.

Let $S(c+c-1)$ be the proper and complete segment set of size $(c+c-1)$, which is formed from $S(c)$ by extending the most probable segment in $S(c)$. Let $S(c+2(c-1))$ be the proper and complete word set of size $(c+2(c-1))$, which is formed from $S(c+c-1)$ by extending the most probable segment in $S(c+c-1)$. Continue this procedure, forming $S(c+k(c-1))$ by extending the most probable segment in $S(c+(k-1)(c-1))$, for $k = 1, 2, \ldots$

Jelinek and Schneider [72] proved that the coding rate converges to the source entropy $H(Q)$ for increasing $T$, i.e.,
\[ \lim_{T \to \infty} R(T) = H(Q) \quad (5.6) \]

**Example 1:**

Given \( X = \{a,b,c,d\} \) with \( Q(a) = 0.6, \ Q(b) = 0.3, \ Q(c) = 0.05, \ Q(d) = 0.05 \). The entropy of this source \( H(Q) \) equals \( 1.40 \) bit/letter. Applying the extension rule of the Tunstall algorithms 4 times, then the size of the segment set is 16. Table 5.1 lists the segments of this set with their corresponding probabilities. \( L = 2.48 \) source-letters/segment. With \( n = 4 \) the coding rate \( R = 1.62 \) bits/letter; this is \( 15.8 \% \) above the source entropy. First-order Huffman coding applied on the same source yields \( R = 1.5 \) bit/letter.

Fig. 5.1 Example of construction of Tunstall code

### 5.2.2 Performance comparison with Huffman codes

The coding rate of the Tunstall code decreases with increasing set size \( T \) and converges to the source entropy as \( T \) approaches infinity (Jellinek & Schneider [72]). The coding rate of first-order Huffman is greater or equal to the source entropy. Hence for every source there exists a Tunstall code which has a lower or equal coding rate than the first-order Huffman code. When constructing the Tunstall code tree, every extension (algorithm step) results in increase of the average segment length and consequently a decrease of the coding rate. We will now derive a bound for the increase of \( L \).

Assume that the segment set \( S_k \) is formed after \( k \) extensions (steps). In this set the string with the largest probability = \( s_k^* \) with probability \( P(s_k^*) \). Assume that \( S_k \) is extended
Table 5.1 Example of VF code (Tunstall algorithm)

<table>
<thead>
<tr>
<th>index</th>
<th>s</th>
<th>L(s)</th>
<th>P(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>aaaa</td>
<td>4</td>
<td>0.1296</td>
</tr>
<tr>
<td>1</td>
<td>aaab</td>
<td>4</td>
<td>0.0648</td>
</tr>
<tr>
<td>2</td>
<td>aaac</td>
<td>4</td>
<td>0.0108</td>
</tr>
<tr>
<td>3</td>
<td>aaad</td>
<td>4</td>
<td>0.0108</td>
</tr>
<tr>
<td>4</td>
<td>aab</td>
<td>3</td>
<td>0.108</td>
</tr>
<tr>
<td>5</td>
<td>aac</td>
<td>3</td>
<td>0.018</td>
</tr>
<tr>
<td>6</td>
<td>aad</td>
<td>3</td>
<td>0.018</td>
</tr>
<tr>
<td>7</td>
<td>ab</td>
<td>2</td>
<td>0.18</td>
</tr>
<tr>
<td>8</td>
<td>ac</td>
<td>2</td>
<td>0.03</td>
</tr>
<tr>
<td>9</td>
<td>ad</td>
<td>2</td>
<td>0.03</td>
</tr>
<tr>
<td>10</td>
<td>ba</td>
<td>2</td>
<td>0.18</td>
</tr>
<tr>
<td>11</td>
<td>bb</td>
<td>2</td>
<td>0.09</td>
</tr>
<tr>
<td>12</td>
<td>bc</td>
<td>2</td>
<td>0.015</td>
</tr>
<tr>
<td>13</td>
<td>bd</td>
<td>2</td>
<td>0.015</td>
</tr>
<tr>
<td>14</td>
<td>c</td>
<td>1</td>
<td>0.05</td>
</tr>
<tr>
<td>15</td>
<td>d</td>
<td>1</td>
<td>0.05</td>
</tr>
</tbody>
</table>

To $S_{k+1}$ by the addition of $c-1$ codewords. This means that segment $s_k^*$ is replaced by $c$ new segments with length $L(s_k^*)+1$. It follows that

$$
\bar{L}_{k+1} = \bar{L}_k - P(s_k^*)L(s_k^*) + \sum_{s \in S_k} P(s_k^*)[L(s_k^*)+1].Q(s)
$$

$$
= \bar{L}_k + P(s_k^*)
$$

$$
\bar{L}_1 = 1 + P(s_0^*) = 1 + Q(0)
$$

(5.7)

$$
\delta_{k+1} = \bar{L}_{k+1} - \bar{L}_k = P(s_k^*)
$$

(5.8)

Probability $P(s_k^*)$ has the following lower bound:

**Lemma:**

$$
P(s_k^*) \geq \max \left\{ 1/T_k, Q(0)^{k-1} \right\}
$$

(5.9)
Proof:

\[ T_k = c + k(c-1) \]

\[ \sum_{s \in S} P(s) = 1 \quad (5.10) \]

Assume \( P(s_k^*) < 1/T_k \). Hence

\[ \sum_{s \in S} P(s) \leq T_k P(s_k^*) < 1 \quad (5.11) \]

This is in contradiction with (5.10). Hence \( P(s_k^*) \geq 1/T_k \). If set \( S_k \) is extended to set \( S_{k+1} \), then a segment with probability \( P = Q(0)P(s_k^*) \) is member of \( S_{k+1} \). Hence \( P(s_{k+1}^*) \geq Q(0)P(s_k^*) \). With \( P(s_0^*) = Q(0) \), it follows that \( P(s_k^*) \geq Q(0)^{k+1} \), Q.E.D.

With a given value of \( T \) there are \( k = \lfloor (T-c)/(c-1) \rfloor \) extensions possible. Summing \( \delta_k \) over all extensions, yields a lower bound on the average segment length

\[ \bar{L}(S_2) \geq 1 + \sum_{i=1}^{k} \text{Max} \left\{ \frac{1}{c+(i-1)(c-1)}, Q(0)^i \right\} \quad (5.12) \]

Substituting \( k \) by \( \lfloor (T-c)/(c-1) \rfloor \) and applying (5.5) yields an upper bound for the coding rate \( R \). This value depends on the properties of the source by the parameters \( c \) and \( Q(0) \). Hence sources with a limited number of source-letters and a large \( Q(0) \) result in low coding rates.

Fig. 5.2 depicts \( R \) for \( T = 256 \), as a function of \( Q(0) \), with \( c \) as a parameter. Obviously, in contrast to first-order Huffman codes, \( R \) becomes less than 1 bit/letter for low-entropy sources.

Being interested in encoding of raster images and line drawings, we will consider a class of DMS sources with the following properties:

a) The probability distribution \( Q(x) \) resembles a double-sided exponential distribution

\[ p(x) = \frac{1}{\sqrt{2\sigma_x}} e^{-\sqrt{2|x|/\sigma_x}} \quad (5.13) \]

which is also known as the Laplacian distribution. Studies show that differential signals corresponding to different images roughly have the same shape and only differ in the variance parameter \( \sigma_x \) (Rabbani [91]). Typical examples of this class of sources are DPCM coded raster images, with \( c \) quantizing levels. Differential Chain Coded (DCC) line drawings also approximate these properties (Liu [92]).
Fig. 5.2 Upper bounds for coding rates of Tunstall Coding, as function of $Q(0)$, for $c = 8, 16$ and $32$ source-letters ($T = 256$ segments)

b) The number of source-letters $c \ll T$. For implementation reasons, we limit the segment set size $T$ to 256, implying that the lexicographical index can be represented by one byte of 8 bits.

Constructing the Tunstall code for a particular source yields a segment set, enabling the calculation of the average segment length $\bar{L}$ according to (5.4), and the coding rate $R$ (5.5). The coding efficiency is defined as the ratio of the source entropy $H(Q)$ and $R$. Fig. 5.3 depicts the calculated coding rates of Tunstall codes ($T=256$, $c=15$) constructed for Laplacian sources, as compared to Huffman codes. As can be expected, the Tunstall codes give a better coding rate performance for small values of $\sigma_x$.

FV coding schemes like Huffman coding, are suffering from word synchronization loss in the presence of random transmission errors. In practical implementations of FV coding, synchronization loss times can be limited by repeated insertion of unique sync patterns. However, this reduces the effective coding rate and increases the complexity. With VF coding all the output codewords have the same length. Hence no word synchronisation loss occurs, resulting in significant lower transmission error sensitivity (Liu [92]).
Fig. 5.3 Coding rates of Tunstall codes as compared to Huffman codes for Laplacian sources, with variance $\sigma_x$ ($c=15$ source-letters, $T=256$ segments)

5.2.3 Software implementation aspects

In this paragraph we will compare fast FV (Huffman) and fast VF (Tunstall) coding schemes on implementation aspects. We consider software-based encoding and decoding algorithms, which are designed to optimise encoding (decoding) speed. It will be shown that VF schemes perform slightly better than FV schemes with respect to speed, but that fast FV requires significantly more memory space than VF.

5.2.3.1 Fast Huffman coding implementations

Tanaka [87] showed that fast Huffman coding can be implemented based on the finite-state machine concept. See also Sieminsky [88]. In this way single-bit manipulation is avoided, and the coder produces fixed-length codewords. Here we will consider 8-bit codewords ($n = 8$). We derived an encoding and a decoding algorithm according to the finite-state machine concept, presented in Tables 5.2 and 5.3 in pseudo code.

The encoder algorithm uses a three dimensional array $M[x,m,i]$, giving the number of fixed-length codewords to be output (0, 1 or 2), the next state and the two codewords, as a function of input letter $x_k$ and previous state $m_k$. It is assumed that the maximum Huffman codeword length is 16 bits, which applies for all sources with $c \leq 16$. 
5.2 Tunstall codes

Table 5.2 Fast Huffman encoder algorithm. Operates on input source-letter stream \( x_k \) and produces output codeword stream \( y_j \).

\[
d := M[x_k, m, 1] \\
\text{if } d > 0 \text{ then} \\
y_j := M[x_k, m, 2] \\
j := j + 1 \\
\text{if } d > 1 \text{ then} \\
y_j := M[x_k, m, 3] \\
j := j + 1 \\
m := M[x_k, m, 4] \\
k := k + 1
\]

Table 5.3 Fast Huffman decoding algorithm. Generates output source-letter stream \( x_j \), operates on input codeword stream \( y_k \).

\[
\text{if } p > 0 \text{ then} \\
x_j := D_2[\text{address } + p] \\
p := p - 1 \\
\text{else} \\
\text{address } := D_1[m, y_k, 1] \\
m := D_1[m, y_k, 2] \\
p := D_2[\text{address}] \\
x_j := D_2[\text{address } + p] \\
p := p - 1 \\
k := k + 1 \\
j := j + 1
\]

The average number of operations per source-letter is equal to

\[
N_{HE} = 4 + 3/\bar{L}
\] (5.14)

\( \bar{L} \) is the average number of source letters which result in the generation of one codeword. The number of possible states depends on the code tree structure. From Table 5.2 it is easily verified that the number of states equals the number of different suffixes of all variable-length codewords, increased by one (no suffix). We calculate this figure for those classes of Laplacian sources which have a \( \sigma_x \) value small enough to result in the maximum number of different codeword lengths: the length of a code assigned to \( Q(j) \) equals \( j + 1 \), except \( Q(c-1) \) of which the length equals \( c-1 \). The number of different suffixes is equal to the maximum
Table 5.4 Example of a Huffman code with a maximum number of different codeword lengths

<table>
<thead>
<tr>
<th>Source-letter</th>
<th>Probability</th>
<th>Codeword</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0.8</td>
<td>0</td>
</tr>
<tr>
<td>b</td>
<td>0.1</td>
<td>10</td>
</tr>
<tr>
<td>c</td>
<td>0.05</td>
<td>110</td>
</tr>
<tr>
<td>d</td>
<td>0.03</td>
<td>1110</td>
</tr>
<tr>
<td>e</td>
<td>0.02</td>
<td>1111</td>
</tr>
</tbody>
</table>

codeword length minus 1 (compare the example of Table 5.4). The number of different suffixes \( D \) equals \( c-2 \). In other words, the state represents the remaining length of a variable-length codeword, which should be incorporated into the next code byte. Hence the number of states equals \( D+1 = c-1 \). However if \( c-2 > n \) then the remaining length is limited to \( n-1 \). This means that for \( c > n+2 \) the number of states is \( n \). Hence the two-dimensional array requires a memory space of

\[
M_{HE} = 4nc \quad \text{bytes} \quad (5.15)
\]

The decoder algorithm uses a two-dimensional array \( D_1[y,m] \), where columns are addressed by the value of a received codeword \( y \) and rows by the decoder state \( m \). The array provides the new state and an address to a segment table \( D_2 \). The latter provides the length of the segment and the values of the corresponding source-letters. The average number of operations per source-letter is equal to

\[
N_{DE} = 4 + 4/\bar{L} \quad (5.16)
\]

The memory occupation of a fast Huffman decoder follows from the analysis of Sieminsky [88]

\[
M_{HD} = (c-1)2^n(3+\bar{L}) \quad \text{bytes} \quad (5.17)
\]

Hirschberg & Lelewer [90] recognized the large amount of memory needed by Sieminsky’s solution and proposed a number of alternatives, however with a reduced performance with respect to decoding time.

The total amount of memory required by a fast Huffman codec is

\[
M_H = 4cn + (c-1)2^n(3+\bar{L}) \quad \text{bytes} \quad (5.18)
\]
5.2.3.2 Fast Tunstall implementations

Table 5.5 represents a high-speed Tunstall algorithm. The two-dimensional array $M_1[x,m]$ provides new encoder state values as a function of the input $x_k$ and the previous state. A zero state indicates the end of a segment and then, by consulting array $M_2[x,m]$, the corresponding codeword can be found. From the algorithm of Table 5.5, the average number of operations per input letter can be derived as

$$N_{TE} = 3 + 2/\bar{L}$$  \hspace{1cm} (5.19)

Table 5.5 Fast Tunstall encoder algorithm. Operates on input source-letter stream $x_k$ and produces output codeword stream $y_j$

<table>
<thead>
<tr>
<th>Condition</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1[x_k,m] = 0$</td>
<td>$y_j := M_2[x_k,m]$</td>
</tr>
<tr>
<td>$m := 0$</td>
<td>$m = \text{state}$</td>
</tr>
<tr>
<td>$j := j + 1$</td>
<td>$M_1 = \text{state table}$</td>
</tr>
<tr>
<td>else</td>
<td>$M_2 = \text{codeword table}$</td>
</tr>
<tr>
<td>$m := M_1[x_k,m]$</td>
<td>$k := k+1$</td>
</tr>
</tbody>
</table>

The number of encoder states is equal to the number of non-terminal nodes in the code tree + 1. In other words, $M_1$ and $M_2$ require

$$M_{TE} = 2c((T-c)/(c-1))^{l+1} \hspace{1cm} (5.20)$$

$$= 2c((T-c)/(c-1))^{l+1} \hspace{1cm} \text{bytes}$$

Table 5.6 presents the decoder algorithm, using a straightforward two-stage look-up data structure. The number of operations per letter can be expressed as

$$N_{TD} = 4 + 3/\bar{L}$$  \hspace{1cm} (5.21)

The decoder tables $D_1$ and $D_2$ require $T$ words (for the addresses into the second table) and $T(1+\bar{L})$ bytes

$$M_{TD} = 2T + T(1+\bar{L}) = 2^n(3+\bar{L}) \hspace{1cm} \text{bytes}$$  \hspace{1cm} (5.22)
Table 5.6  Fast Tunstall decoding algorithm. Generates output source-letter stream $x_j$; operates on input codeword stream $y_k$

<table>
<thead>
<tr>
<th>if $l &gt; 0$ then</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_j := D_2[address + l]$</td>
</tr>
<tr>
<td>$l' := l - 1$</td>
</tr>
<tr>
<td>else</td>
</tr>
<tr>
<td>$k := k + 1$</td>
</tr>
<tr>
<td>address := $D_1[y_k]$</td>
</tr>
<tr>
<td>$l := D_2[address]$</td>
</tr>
<tr>
<td>$x_j := D_2[address + l]$</td>
</tr>
<tr>
<td>$l' := l-1$</td>
</tr>
<tr>
<td>$j := j + 1$</td>
</tr>
</tbody>
</table>

From (5.20) and (5.22) the total memory requirements of the Tunstall codec can be derived

$$M_T = 2\epsilon((2^n-c)/(c-1))+1) + 2^n(3-L) \text{ bytes}$$ (5.23)

The Tunstall algorithms appear to be slightly faster than the Huffman equivalents (compare (5.14) and (5.16) with (5.19) and (5.21)). Table 5.7 lists $M_T$ and $M_H$ for typical applications ($L=2$, $n=8$) and various values of $c$. Tunstall codecs requires less than 2 kByte, while fast Huffman implementations requires between 2 and 23 times more memory space.

Table 5.7  Comparison of memory occupation of Fast Huffman and Tunstall codecs ($L=2$, $n=8$), in bytes

<table>
<thead>
<tr>
<th>$c$</th>
<th>Huffman</th>
<th>Tunstall</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3968</td>
<td>1960</td>
</tr>
<tr>
<td>8</td>
<td>9216</td>
<td>1856</td>
</tr>
<tr>
<td>16</td>
<td>19712</td>
<td>1824</td>
</tr>
<tr>
<td>32</td>
<td>40704</td>
<td>1792</td>
</tr>
</tbody>
</table>
5.3 Modified Tunstall Coding (MTC)

5.3.1 The MTC algorithm

During the construction of the Tunstall code tree, every extension taking \( c-1 \) codewords of the total 'budget' \( T \), results in an increase of the average segment length. The total number of possible extensions is limited to \( k_T = \lceil (T-c)/(c-1) \rceil \). The larger \( T \), the larger \( k_T \) becomes and correspondingly the closer the source entropy is approached.

In this section we propose a modified version of the Tunstall algorithm, with a larger number of extensions within a given budget, and, consequently, generally a coding rate closer to the source entropy. This type of VF coding, abbreviated as MTC, applies a merge-and-split rule on the source alphabet. At the encoder the merge operation implies that all values of \( x_k \geq F-1 \) are replaced by the value \( F-1 \). On the reduced alphabet of \( F \) source-letters, normal Tunstall coding is applied, however with an adaptation of the construction algorithm guaranteeing a coded segment containing not more than one source-letter of the value \( F-1 \). In order to enable the decoder to reconstruct \( x_k \), each codeword representing a segment with a source-letter \( F-1 \) (the lexicographical index) is followed by a codeword with a value equal to the original \( x_k \). In this way, according to (5.1), the \( c-F+1 \) source-letters with the smallest probabilities are combined. The parameter \( F \) will be referred to as the merge delimiter. Let

\[
Q_T = \sum_{i=F-1}^{c-1} Q(x_i)
\]  

(5.24)

For each source there exists an optimum of \( F \). Although small values of \( F \) will result in a high number of extensions, resulting in a Tunstall code with a coding rate close to the source entropy, the probability of additional codewords is increased. The Tunstall code tree construction algorithm is modified as follows:

Algorithm 2:
Let \( S(F) = X_F \). This is the complete and proper segment set of size \( F \). Let \( S(F+F-1) \) the proper and complete segment set of size \( (F+F-1) \), which is formed from \( S(F) \) by extending the most probable segment in \( S(F) \) except segments ending at source-letter \( F-1 \). Let \( S(F+2(F-1)) \) be the proper and
complete word set of size \( (F+2(F-1)) \), which is formed from \( S(F+F-1) \) by extending the most probable segment in \( S(F+F-1) \) except segments ending at source-letter \( F-1 \). Continue this procedure, forming \( S(F+k(F-1)) \) by extending the most probable segment in \( S(F+(k-1)(F-1)) \), for \( k = 1, 2 \ldots \)

The following example clarifies this modified approach.

**Example 2:**
Given \( X = \{a,b,c,d,e\} \) with \( Q(a) = 0.6, Q(b) = 0.3, Q(c) = 0.05, Q(d) = 0.03, Q(e) = 0.02 \). Assume that on this source a modified Tunstall algorithm is applied with \( c=5 \) \( (F = 4) \). The algorithm converts all source-letters with value \( e \) into value \( d \). For \( T=16 \) the optimal segment set is equal to the one of example 1. When a codeword is produced representing a segment with the value \( d \), the codeword is followed by a second codeword with value \( d \) or \( e \), dependent on the original source letter value.

Assuming an input stream

\[
x = \{a,a,a,a,b,a,a,c,a,d,a,a,b,a,e,a,a,a,b\},
\]

then the coder produces the codeword stream

\[
y = \{0,10,8,9,A,4,9,5,1\}
\]

The bold values represent segments with a letter from the set \( \{F-1, \ldots, c-1\} \). Therefore these values are followed by the value of the original source-letter; the latter are underlined.

We will now derive an expression for the gain in coding efficiency of MTC over TC, for a given value of \( T \). Let a TC code tree have \( k_T+1 \) non-terminal nodes: the initial node increased with \( k_T \) extensions. If the source has \( c \) letters, then \( k_T \) is expressed as

\[
k_T = \lfloor T-c/c-1 \rfloor \tag{5.25}
\]

Let a MTC code tree be constructed with the same source, and have \( k_M+1 \) non-terminal nodes, or \( k_M \) extensions

\[
k_M = \lfloor T-F/F-1 \rfloor \tag{5.26}
\]

The average MTC segment length \( \bar{L}_M \) will be greater than the average TC segment length \( \bar{L}_T \), because \( k_M > k_T \).
The following expressions can be derived from (5.7), taking into account that $\bar{L}_T = 1 + P(s_0^*) = 1 + Q(0)$

$$\bar{L}_T = 1 + \sum_{k=0}^{k_T-1} P(s_k^*)$$  \hspace{1cm} (5.27)

$$\bar{L}_M = 1 + \sum_{k=0}^{k_M-1} P(s_k^*)$$  \hspace{1cm} (5.28)

$$R_T = \frac{[\log_2 T]}{\bar{L}_T}$$  \hspace{1cm} (5.29)

$$R_M = \frac{(1+p_s)\log_2 T}{\bar{L}_M}$$  \hspace{1cm} (5.30)

Here $p_s$ denotes the probability of a second codeword. Every extension will add a segment ending with a 'F-1 letter'. Therefore

$$p_s = \sum_{k=0}^{k_M-1} P(s_k^*) Q_L = Q_L(\bar{L}_M-1)$$  \hspace{1cm} (5.31)

Hence the gain in coding efficiency is expressed as

$$G_{MTC} = \frac{R_T}{R_M} = \frac{\bar{L}_M}{(1-Q_L+Q_L\bar{L}_M)\bar{L}_T}$$  \hspace{1cm} (5.32)

Combination of (5.27), (5.28) and (5.31) gives an expression for the average segment length

$$\bar{L}_M = \bar{L}_T + \sum_{k=k_T}^{k_M-1} P(s_k^*) = \bar{L}_T + \beta_M$$  \hspace{1cm} (5.33)

and for the gain in coding efficiency of MTC over TC
\[ G_{MTC} = \frac{\bar{L}_T + \beta_M}{(1 - Q_L + Q_L\bar{L}_T + Q_L\beta_M)\bar{L}_T} \]
\[ = \frac{\bar{L}_T + \beta_M}{\bar{L}_T} \cdot \frac{1}{1 + Q_L(\bar{L}_T + \beta_M - 1)} \]

(5.34)

From this expression it follows that there exists an optimum value for \( F \), since decreasing \( F \) will result in a larger \( k_M \) and consequently a larger \( \beta_M \). However \( \beta_M \) is limited by the source entropy. On the other hand decreasing \( F \) will result in an increase of \( Q_L \), decreasing the second term. To ensure that \( G_{MTC} > 1 \), it should be fulfilled that

\[ \beta_M > Q_L(\bar{L}_T + \beta_M - 1) \quad \Rightarrow \quad \frac{\beta_M}{\bar{L}_T} > \frac{Q_L(\bar{L}_T - 1)}{1 - Q_L\bar{L}_T} \]

(5.35)

The next section evaluates \( G_{MTC} \) as a function of \( F \) for a given source and \( T \).

### 5.3.2 Performance of Modified Tunstall Coding

Figure 5.4 illustrates that the coding rate of MTC converges faster than TC to the source entropy. The coding efficiency is plotted as a function of the segment set size \( T \) and has been calculated for a Laplacian source with \( \sigma_x = 0.5 \), and \( F = 6 \). The MTC coding efficiency saturates above \( T = 100 \), due to the increasing probability of a second codeword. Although this result could lead to the conclusion that a set size \( T = 128 \) would be preferable, one should realize that 8-bit codewords are necessary in order to realize fast software algorithms. Fig. 5.5 depicts the MTC coding efficiency as a function of \( F \) for two examples: \( \sigma_x = 0.5 \) and \( \sigma_x = 0.8 \). As expected (see 5.3.1), the results show that there exists an optimum \( F \) value, which depends on the source statistics. Finally, Fig. 5.6 compares MTC coding efficiency (with optimum \( F \) values) to TC and Huffman as calculated for various Laplacian sources (function of \( \sigma_x \)). The conclusion is that MTC performs better than TC over the complete range and even better than Huffman for \( \sigma_x < 0.8 \).
Fig. 5.4 Comparison of convergence rates for MTC (merge delimiter $F = 6$) and TC (Laplacian source, with variance $\sigma_x = 0.5$)

5.3.3 Implementation aspects of MTC

Table 5.8 resembles an encoder algorithm optimized with respect to speed. For the decoder the same algorithm as for TC can be applied (Table 5.6). The encoder algorithm uses a finite-state automaton concept, enabling look-up table techniques. Every source letter $x_k$ results in a new state value $m_{k+1}$, determined by a $c \times T$ two-dimensional array $M[x,m]$. If $m_{k+1}$ appears to be zero, the previous letter(s) represent a completed segment, which is coded with a value equal to the previous state $m_k$. Note that this approach combines the Tunstall coding algorithm with the split operation into one finite-state automaton, resulting in the fastest possible implementation. The difference with the implementation of the normal Tunstall encoder is twofold: i) The initial code tree contains all $c$ source-letters (while the others branches only contain $F$ letters); ii) The coding is delayed by one source-letter. The decoder is simply realized with a two-stage look-up data structure, similar with Table 5.6.

It is easily verified that this MTC method equals the performance of TC with respect to speed. The memory space requirements of the decoder are also the same. However the encoder requires more memory.
Fig. 5.5 Coding efficiency of MTC for $\sigma_x = 0.5$ and $\sigma_x = 0.8$ as a function of merge delimiter $F$

Table 5.8 Fast Modified Tunstall encoder algorithm. Operates on input source-letter stream $x_k$, and produces output codeword stream $y_j$; initially $m=0$

```
if M[x_k,m] = 0 then
    j := j + 1
    y_j := m
    m := x_k + 1
else
    m := M[x_k,m]
    k := k + 1
```

The number of states equals the number of codewords

$$M_{ME} = cT$$  \hspace{1cm} (5.36)

For a typical application $M_{ME}$ becomes 256 times 16 words = 4096 bytes.
Fig. 5.6 Coding efficiency of MTC compared to TC and Huffman for Laplacian sources as a function of variance $\sigma_X$

Example 3:

Fig. 5.7 shows the code tree for a source with $c = 4$ ($Q(a)=0.7$, $Q(b)=0.2$, $Q(c)=0.05$, $Q(d)=0.05$), $F = 3$, $T = 6$. Note that the initial tree contains all $c$ source-letters, in order to combine the split operation and the Tunstall algorithm in one finite-state automaton. Table 5.9 represents the corresponding encoder array. Fig. 5.8 gives the decoder data structures.

Fig. 5.7 MTC code tree for $c = 4$ letters, merge delimiter $F = 2$, number of segments $T = 8$
Table 5.9 MTC encoder finite state automaton example corresponding to code tree of Fig. 5.7

<table>
<thead>
<tr>
<th>m</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Fig. 5.8 MTC decoder data structures corresponding to the code tree of Fig. 5.7

5.3.4 Sensitivity of MTC to varying source statistics

In practical situations, sources are seldom characterized by stationary statistics. Therefore it is important to investigate whether the performance of a coding method is sensitive for these variations. In Fig. 5.9 and 5.10 the coding efficiencies of an MTC code, a TC code and a Huffman code are plotted as function of $\sigma_x$, which were optimized for $\sigma_x = 0.5$ and $\sigma_x = 1.0$, respectively.
Fig. 5.9 Coding efficiency of MTC, TC and Huffman coding (optimized for \( \sigma_x = 0.5 \)) as a function of variance \( \sigma_x \)

MTC can be adapted to the source statistics by selection of the appropriate code using a sample of the \( Q(0)/Q(1) \) ratio. From (5.13) it is clear that this relation depends on \( \sigma_x \), since

\[
\frac{Q(0)}{Q(1)} = e^{\sqrt{2\sigma_x}} \tag{5.37}
\]

Fig. 5.11 compares the results of Adaptive MTC (AMTC) with TC and Huffman. In this case three MTC codes were applied: on the interval \( 0.3 \leq \sigma_x \leq 0.5 \) the MTC optimized for \( \sigma_x = 0.4 \); on the interval \( 0.6 \leq \sigma_x \leq 0.8 \) the MTC optimized for \( \sigma_x = 0.7 \); on the interval \( 0.9 \leq \sigma_x \leq 1.2 \) the MTC optimized for \( \sigma_x = 1.0 \). Apparently AMTC results in a high coding efficiency level over a wide \( \sigma_x \) range.
Fig. 5.10 Coding efficiency of MTC, TC and Huffman coding (optimized for $\sigma_x = 1.0$) as a function of variance $\sigma_x$.

Fig. 5.11 AMTC compared to TC and Huffman coding (the latter two codes are optimized for $\sigma_x = 0.8$)
5.4 Summary and conclusions

High-speed Huffman coding, implemented in software, requires large amounts of memory. It has been shown that Tunstall coding can be applied with the same or better performance with respect to the coding rate (depending on the source statistics), and with significant lower complexity. Moreover, variable-to-fixed-length coding schemes like Tunstall’s algorithm, are less sensitive to transmission errors.

A modified version of the Tunstall coding method has been proposed (MTC), with a coding rate closer to the source entropy due to a faster convergence rate. For MTC, complexity aspects were discussed, as well as the sensitivity to variations of the source statistics.

In conclusion, MTC is an attractive choice for software-implemented entropy coding in the case of limited alphabet sources with a skew probability distribution.
References in Chapter 5


Huffman, D.A. [52], "A method for the construction of minimum redundancy codes", *Proc. IRE*, Vol.40, pp.1098-1101


Tunstall, B.P.[67], "Synthesis of noiseless compression codes", *Ph.D. dissertation*, Georgia Institute of Technology, Atlanta, U.S.A.

6 SOME IMPROVEMENTS OF TELEGRAPHIC CODING APPLIED TO REAL-TIME TELE-EDUCATION

6.1 Introduction

Telewriting can be defined as the transmission of handwritten information produced by an electronic writing tablet by means of a hand-held stylus. The increasing use of electronic mail, remote databases and multimedia communication systems require methods for representing graphical information which cannot be covered by the standard ASCII character set. Examples of applications are: mathematical formulas, music scores, Chinese and Japanese characters, handwritten annotations, signatures. A robust but efficient method for encoding handwritten information has been proposed in Kegel & Bons [77] and Bordewijk [78], called Differential Chain Coding (DCC). In Arnbak, Bons & Vieveen [89] the use of DCC in PC-based electronic mail systems is described.

Telewriting is a useful facility in real-time tele-education systems. It offers the possibility to the teacher to support his verbal explanation by on-line handwritten graphics, much like the blackboard in face-to-face education. Some typical examples of systems where some form of telewriting has been implemented are: UWIDITE (University of the West-Indies Distance Teaching Experiment) (Lalor [84]), CYCLOPS (Great Britain’s Open University distance teaching system) (Mc.Connell [86]), and the ISEMSI (Isolation Study European Manned Space Infrastructure) tele-instruction experiment for ground-to-astronaut communication (Novara, Collet & Vaernes [91], Bruyne [91]).
The application of telewriting in real-time tele-education systems implies a number of specific requirements and constraints:

1) Real-time encoding and transmission of telewriting information should be performed without noticeable delay, in order to maintain a convincing relation with the speech signal in the teaching process.

2) An encoding method should be selected which provides effective compression, as well as robustness against transmission bit errors.

3) The encoding method should be appropriate for software implementation (i.e., be computationally and memory efficient), since hardware (de)compression systems are not commonly available for this type of application.

4) The use of telewriting in a real-time multimedia environment demands efficient user-interface functionality

This chapter proposes a number of improvements of telegraphic coding which support these requirements. Section 6.2 deals with two functional system enhancements. First, a colour fill method in combination with telewriting, which is very efficient with respect to coding efficiency and which allows effective object-oriented editing. Second, an efficient solution is given to handle the eye-to-hand coordination problem, which is inherent in the use of a writing tablet as a PC-peripheral device. In section 6.3 it is proved that DCC is a particular, but sub-optimal form of Modified Tunstall Coding (MTC). It is shown that with MTC a significant reduction of the coding rate is possible, without losing DCC's advantage of robustness against transmission errors.
6.2 Functional enhancements

6.2.1 Colour-fill facility

In this section, Differential Chain Coding is briefly described and an extension is proposed, constituting an efficient colour-fill facility.

6.2.1.1 Differential Chain Coding

The basic underlying principle of DCC is spatial coding, which means that a fixed amount of data is generated for any given line segment (Arnbak, Bons & Vieveen [89]). The trace coordinates are represented as a chain of vectors on the asynchronous, time-independent functional form

$$
\begin{align*}
  x_{k+1} &= F(x_k, y_k) \\
  y_{k+1} &= G(x_k, y_k)
\end{align*}
$$

(6.1)

Here, $k$ is a sampling-point index which is stepped up (only) when the boundary of an associated ring of reference pixels around the sampling point is crossed by the trace. Where the growing trace intersects the ring, the nearest reference pixel is selected as the following sampling point $(k+1)$ - see Fig. 6.1. It proves advantageous to encode and transmit the differences between successive vectors, i.e., the relative displacements:

$$
\begin{align*}
  \delta x_{k+1} &= x_{k+1} - x_k \\
  \delta y_{k+1} &= y_{k+1} - y_k
\end{align*}
$$

(6.2)

It should be noted that no new data emerge from a DCC encoder until justified by a sufficient displacement of the stylus. Therefore DCC lends itself to graphical applications in which storage capacity or processing power are important constraints, and to asynchronous data networks. Another attractive advantage from a user point of view is the fact that the decoding and display process follows the dynamographical movements of the pen. This causes the subjective impression of a much faster displaying process than with raster-scanning methods (e.g., facsimile). Due to the statistical properties of writing, the probability of large variations between subsequent vector displacements (6.2) is small.
Fig. 6.1 Quantization process using Differential Chain Coding for a square coding ring with N=16

Although fixed-to-variable-length (FV) coding of these vector differences is closer to the information-theoretical optimum, a string-incremental method (Arnbak, Bons & Vieveen [89]) overcomes a number of practical disadvantages of FV due to variations in probability distributions and the sensitivity of FV to transmission errors. Any curve trace between pen-down and pen-up can be encoded as a chain of absolute vector displacements (6.1) and incremental strings. Each of these chain elements fits the format of a byte and therefore it is feasible to represent and communicate dynamographical information about line drawings with a data syntax based on CCITT alphabet no. 5 (i.e., ASCII-characters). The header of such a chain of data contains additional information such as colour and thickness of the line.

6.2.1.2 Colour-fill methods

In a number of teaching and other applications the need arises to fill certain areas on the screen with suitable colours, e.g., in maps. In the case of the DCC method, this is only possible by 'painting'. In other words: Fill each area to be coloured with a number of lines with the appropriate colour and thickness. This method is inconvenient and also requires a lot of additional DCC data; therefore it is not efficient with respect to transmission and storage. Another way to approach this problem is the use of filling algorithms (Foley et al. [90], Shani [80], Pavlidis [81]). A region to be coloured can be defined in two ways. For either we use a starting pixel, P. The interior-defined region is the largest connected region of points whose value is the same as that of P. A boundary-defined
**region** is the largest connected region of pixels whose value is not some given boundary value. Algorithms that fill interior-defined regions are called *flood-fill algorithms*; those that fill boundary-defined regions are called *boundary-fill algorithms* (Foley et al. [90]). Although very efficient from a coding point of view, these techniques have some major drawbacks:

1) Editing of screen objects can change the boundary conditions and therefore cause unexpected and uncontrolled changes in the colour-filled areas. Thus, editing of screens containing colour-filled objects is hindered.

2) In advanced graphical applications, world-coordinates systems are used, enabling the use of fast algorithms which can perform geographical changes like translation, rotation, zooming, and window clipping in a screen-independent way. Since most of the filling algorithms are related to the actual screen contents, its use in combination with world-coordinates is complicated.

3) The region to be filled must be truly closed. In other words, the boundary must be complete and unbroken, all the way around. Even a 1-pixel gap will let the fill pattern or colour 'leak out' and spread to fill the entire remaining empty area of the screen.

An approach for a colour-fill facility without these disadvantages, is the following procedure. When filling is to be applied, three steps have to be taken (see Fig. 6.2):
Procedure
1) Define a screen-point \((x_F, y_F)\), which will act as a centre for the filling operation.
2) Define the fill colour
3) Draw a trace (with separately defined colour and thickness) as a boundary of the fill area.

The filling operation is performed directly as follows: A straight line in the fill colour is drawn between each new sampling point \((x_{k+1}, y_{k+1})\) along the trajectory and the centre point \((x_F, y_F)\). The proposed method has the following advantages:

a) The boundary and the fill area are generated as one integrated object. This enables performing graphical editing functions without problems.
b) The additional fill command information to be transmitted consists of just three elements: the fill colour and the coordinate pair \((x_F, y_F)\).

Fig. 6.3 shows how the trace header of a Differential Chain Coded trajectory can be extended with this information.

6.2.2 Relative coordinate system (RCS) facility

In some cases, application of handwriting by the use of a hand-held stylus and an electronic writing tablet, raises a eye-to-hand coordination problem. Normally a VDU (Visual Display Unit) is used for immediate presentation and check of the drawn objects. If an existing page displayed on the VDU has to be extended or annotated by handwritten graphics, it is difficult to position the pen correctly. Often, precise positioning relative to existing objects is necessary. Therefore a Relative Coordinate System facility (RCS) should be available (to the user), providing a convenient method to overcome the coordination problem.

We will describe two possible implementations of such an RCS facility, both making use of a cursor on the VDU screen which must be positioned prior to the actual handwriting action. Normally the active tablet region is mapped 1:1 on the complete screen (or the active window), involving similar coordinate systems for tablet and display. However, the RCS function defines a temporary coordinate transformation. The cursor must be positioned on that location \((x_r, y_r)\) where the new object is to be inserted. Let the subsequent drawing activity start at an arbitrary tablet location \((x_t, y_t)\). Then the RCS
Fig. 6.3 Frame structure of DCC with colour-fill enhancement. The fill-mode is activated when the fill-colour byte represents a valid fill-colour. The pen-up code deactivates filling.

function performs a displacement \((x_r-x_f,y_r-y_f)\) on all following pixels generated by the tablet. Method 1 uses the pen of the tablet as a cursor positioner. Method 2 applies another pointing device (e.g., a mouse), independent of the tablet. These methods are written as step-by-step procedures:

Procedure Method 1:

Step 1: Activate RCS function (keystroke).
The cursor appears on the VDU screen.

Step 2: Position the cursor (by moving the pen over the tablet pad)

Step 3: Switch to drawing mode (keystroke)

Step 4: Actual drawing starting at an arbitrary tablet location, corresponding to a cursor-defined screen position

Step 5: Deactivate RCS function (keystroke)

Deactivation restores the original 1:1 coordinate mapping. The cursor disappears.
Procedure Method 2:

Step 1: Activate RCS function (keystroke)
The cursor appears on the VDU screen.
Step 2: Position the cursor (by moving the mouse)
Step 3: Start drawing at an arbitrary tablet location, corresponding to a
cursor-defined screen position.
Step 4: Deactivate RCS function (keystroke or mouse-button). Deactivation
restores the original 1:1 coordinate mapping. The cursor disappears.

Method 2 is preferred because of the following advantages:

a) This method is faster since it requires only 4 steps of action (method 1 requires 5
steps). Moreover, the user can choose to use both hands: one hand for cursor
positioning and the other for drawing, allowing a convenient and rapid way of
working.

b) The method does not result in inadvertent ink stripes on the tablet paper due to pen-
controlled cursor movement

c) Repetition of the function is realised very effectively: as soon as the cursor is
moved to another location, a new transformation is performed. In other words,
intermediate key strokes are not required.

6.3 Variable-to-fixed-length coding

6.3.1 Relation of DCC to Modified Tunstall Coding (MTC)

Fig. 6.1 shows the quantization process of DCC using a square coding ring with \(N=16\)
discretisation points. Let \(\tau\) represent the distance between two adjacent discretisation
points on the ring, defined as a square of side \(2M\tau\) and with \(N =8M\) equispaced
discretisation points. These are arranged such that there is a point at every corner, and
numbered from \(-N/2+1\) to \(N/2\). Any vector directed from the centre of the ring to an
arbitrary discretisation point is an absolute vector \(A_i, i \in \{-N/2+1,...,0,...,N/2\}\). Initially,
the coding ring is centred on the starting point of the line-drawing. The first vector
displacement along the line-drawing is chosen as the best fitting of the \(N\) absolute vectors,
as shown in Fig. 6.1. The end of a vector becomes the centre for the next ring. The
discretisation point nearest to each intersection between the line-drawing and the coding
ring is determined. The increment in the direction of each subsequent absolute vector can
be determined and defines a relative vector

$$D_k = D_{(j-i)}$$

(6.3)

where \(j\) and \(i\) are the index of the corresponding absolute vector \(A_j\) and its predecessor \(A_i\),
and \(k = j-i = -(N-1),\ldots,0,\ldots,(N-1)\). Each step along the curve is now encoded as one of the
relative vectors if \(-R \leq k \leq R\), where the parameter \(R\) is referred to as the relative
vector delimiter. Otherwise a step is encoded as the appropriate absolute vector. This
coding principle is illustrated in Fig. 6.1 for \(R=1\), in which the vectors 2,3,4 and 5 are
all encodable as a relative vector, while vector No. 1 (by definition) and No. 6 (by
deviating -2, relative to its predecessor) will be encoded as absolute vectors. In this way,
yany line drawing can be encoded as a chain of absolute and relative vectors. Smooth
curves will result in a high percentage of relative vectors (Arnbak, Bons & Viveen [89],
Prasad et al.[89]). More efficient encoding of chains of absolute and relative vectors can
be achieved by assigning codewords of length \(b_r\) to the relative vectors and of length \(b_a\)
to the absolute vectors, where \(b_r < b_a\). For one-bit-prefix DCC (Liu [92]) this results in
an optimum for: \(b_r = 1 + \lceil \log_2(2R+1) \rceil\) and \(b_a = 1 + \lceil \log_2 N \rceil\).

In Arnbak, Bons & Viveen [89] a practical implementation is described for \(R=1\),
based on fixed-length coding (using codewords with a length of 6 bits). The input vector
stream is chopped into segments of subsequent vectors chosen from a segment set as
defined in Table 6.1 (\(N=8\)). The segment set resembles a complete but not proper
ensemble of vector segments (Jelinek & Schneider [72]). However, encoding implies that
the longest segment possible is always used. To every segment a fixed-length codeword is
assigned. The maximum segment length is 3. The size of the segment set equals

$$T_{DCC} = N + \sum_{i=1}^{L} (2R+1)^i$$

(6.4)

and the coding rate is as follows (Liu [92])

$$B_{DCC} = \frac{(1+p_r^L)(1-p_r^L)}{1-p_r^L} \lceil \log_2 T_{DCC} \rceil$$

(6.5)

where \(p_r\) denotes the probability of a relative vector.
Table 6.1 Segment set of DCC (N=8, R=1). The relative vectors $D_0$, $D_1$ and $D_{-1}$ are denoted by '0', '+' and '−', respectively.

<table>
<thead>
<tr>
<th>Code-word</th>
<th>Segment</th>
<th>Code-word</th>
<th>Segment</th>
<th>Code-word</th>
<th>Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$A_0$</td>
<td>16</td>
<td>+−</td>
<td>32</td>
<td>++0</td>
</tr>
<tr>
<td>1</td>
<td>$A_1$</td>
<td>17</td>
<td>−0</td>
<td>33</td>
<td>++0</td>
</tr>
<tr>
<td>2</td>
<td>$A_{-1}$</td>
<td>18</td>
<td>−+</td>
<td>34</td>
<td>++−</td>
</tr>
<tr>
<td>3</td>
<td>$A_2$</td>
<td>19</td>
<td>−−</td>
<td>35</td>
<td>+−0</td>
</tr>
<tr>
<td>4</td>
<td>$A_{-2}$</td>
<td>20</td>
<td>000</td>
<td>36</td>
<td>+−+</td>
</tr>
<tr>
<td>5</td>
<td>$A_3$</td>
<td>21</td>
<td>00+</td>
<td>37</td>
<td>+−−</td>
</tr>
<tr>
<td>6</td>
<td>$A_{-3}$</td>
<td>22</td>
<td>00−</td>
<td>38</td>
<td>−00</td>
</tr>
<tr>
<td>7</td>
<td>$A_4$</td>
<td>23</td>
<td>0+0</td>
<td>39</td>
<td>−0+</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>24</td>
<td>0+ +</td>
<td>40</td>
<td>−0−</td>
</tr>
<tr>
<td>9</td>
<td>+</td>
<td>25</td>
<td>0+−</td>
<td>41</td>
<td>−+0</td>
</tr>
<tr>
<td>10</td>
<td>−</td>
<td>26</td>
<td>0−0</td>
<td>42</td>
<td>−++</td>
</tr>
<tr>
<td>11</td>
<td>00</td>
<td>27</td>
<td>0−+</td>
<td>43</td>
<td>−−−</td>
</tr>
<tr>
<td>12</td>
<td>0+</td>
<td>28</td>
<td>0−−</td>
<td>44</td>
<td>−−0</td>
</tr>
<tr>
<td>13</td>
<td>0−</td>
<td>29</td>
<td>+00</td>
<td>45</td>
<td>−−+</td>
</tr>
<tr>
<td>14</td>
<td>+0</td>
<td>30</td>
<td>+0+</td>
<td>46</td>
<td>−−</td>
</tr>
<tr>
<td>15</td>
<td>++</td>
<td>31</td>
<td>+0−</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Now it will be shown that the described form of DCC is a particular case of Modified Tunstall Coding (MTC), the variable-to-fixed-length coding method as presented in Chapter 5 (see also Algra [92]).

**Lemma:** Differential Chain Coding is a particular case of Modified Tunstall Coding

**Proof:**
Consider a discrete memoryless source (DMS) with source letter alphabet $X_c$ and probability distribution $Q(x)$, $x \in X_c$. $X_c = \{0,1,2,\ldots,c-1\}$. Consider two variables $R$ and $N$ such that $c = 2R+1+N$. Assume a $Q(x)$ fulfilling the following three properties:

1) $Q(0) = Q(1) = \ldots = Q(2R) = p_r/(2R+1)$
2) $Q(2R+1) = Q(2R+2) = \ldots = Q(c-1) = (1-p_r)/N$
3) $p_r/(2R+1) > (1-p_r)/N$

In contrast to MTC, the construction of the DCC segment set is independent of the source statistics. We will show that the MTC segment set as constructed for this particular source is equal to the DCC set. Let the MTC code construction rule be
applied on the source (see Chapter 5) with the *merge delimiter* set to 2R+2. The code construction process implies subsequent extensions of the code tree with 2R+1 new branches (segments). Let the process be continued until the total number of nodes (segments) equals \( T \) as given in eqn. (6.4). According to the code construction rule, each time that terminal node of the code-tree is extended, which represents the segment with the largest probability. The probability of a segment of length \( s \ (s > 1) \) is equal to \((p_s/2R+1)^s\). Thus all the terminal nodes representing a string length \( s \) are extended, before any terminal node representing string length \( s+1 \) is extended. In other words, the eventual segment set contains \((2R+1)^s\) segments of length \( s \), \( s = 1,2,...L \). Hence the resulting segment set is similar to the DCC encoding scheme. *Q.E.D.*

6.3.2 MTC applied to telewriting

It will be shown that DCC is a sub-optimal MTC variant, due to two factors.

1) Assume fixed-length codewords of \( n \) bits. Hence \( T \leq 2^n \). Then, generally, the size of the DCC segment set, \( T_{DCC} \) (eqn. 6.4), will be considerably less than the set size \( T_{MTC} \):

\[
T_{MTC} = k(2R+1)+N \quad k = 1,2,3,...
\]

For example, \( R=1, n=6, N=8 \) results in \( T_{DCC} = 47 \) and \( T_{MTC} = 62 \). Obviously, DCC is characterized by a smaller average segment length, which implies a higher coding rate.

2) Experimentally obtained relative vector probabilities for a set of line-drawings are presented in Table 6.2 (according to Liu [92]). Obviously, the probability distribution of the relative vectors deviates significantly from the uniform model as used in DCC. Hence, when MTC is applied using a source-model with a non-uniform relative vector probability which approximates the experimentally derived statistics, lower coding rates will result.

We compared the coding rate of DCC to MTC for \( M=1,2 \) and 3 with \( R=1 \), for a codeword-length of 6 bits. The source model used, resembled the relative vector probabilities as presented in Table 6.2 and a uniform absolute vector distribution. The latter is a reasonable approximation of the long-term averaged probability density of the angular direction for line drawings, as has been shown in Liu [92].
Table 6.2 Relative vector probabilities for a set of line drawings as quantized using a square coding ring, for $N = 8$, $N = 16$, $N = 24$ (according to Liu [92], p. 100)

<table>
<thead>
<tr>
<th>$D_1$</th>
<th>$N = 8$</th>
<th>$N = 16$</th>
<th>$N = 24$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_0$</td>
<td>0.6778</td>
<td>0.4924</td>
<td>0.3909</td>
</tr>
<tr>
<td>$D_1$</td>
<td>0.1170</td>
<td>0.1437</td>
<td>0.1436</td>
</tr>
<tr>
<td>$D_{-1}$</td>
<td>0.1170</td>
<td>0.1437</td>
<td>0.1436</td>
</tr>
</tbody>
</table>

Table 6.3 Line drawing coding rates for DCC and MTC, in bit/vector

<table>
<thead>
<tr>
<th>$N$</th>
<th>Entropy</th>
<th>$B_{DCC}$</th>
<th>$B_{MTC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1.678</td>
<td>2.524</td>
<td>1.995</td>
</tr>
<tr>
<td>16</td>
<td>2.669</td>
<td>3.281</td>
<td>3.082</td>
</tr>
<tr>
<td>24</td>
<td>3.336</td>
<td>3.834</td>
<td>3.796</td>
</tr>
</tbody>
</table>

Table 6.3 gives the source entropy and the resulting coding rates of DCC and MTC. $B_{DCC}$ was calculated from (6.5). $B_{MTC}$ was determined (see Chapter 5) after construction of the coding table, according to the MTC construction rule. Table 6.4 shows the resulting segment set for $M=1$. The maximal segment-length appears to be 8 vectors, which is considerably more than with DCC. Apparently, MTC performs better than DCC with respect to the coding rate, by up to 30% higher coding efficiency.

An important advantage of DCC as compared to Huffman coding (Huffman [52]) is its resistance to transmission bit-errors. The occurrence of a bit error in Huffman-encoded data can result in the loss of word synchronization at the decoder. Consequently, an unpredictable number of subsequent codewords will be falsely decoded. However with fixed-length coding no loss of word synchronization occurs. The latter applies to DCC as well as MTC. Another property of DCC which contributes to its robustness, is the termination of the propagation effect of a bit error as soon as an absolute vector is received (correctly). Other schemes like Chain Difference Coding (Lu & Dunham [91]), where only differential vectors are encoded, lack such a built-in protection against error propagation. Furthermore DCC lends itself to a simple but effective additional measure against bit errors, as described in Klerk [90]. Such a higher protection level is achieved through a limit on the maximum length allowed for each trace and on the maximum
Table 6.4 Segment set of MTC (N=8, R=1). The relative vectors $D_0$, $D_1$ and $D_{-1}$ are denoted by '0', '+' and '-', respectively.

<table>
<thead>
<tr>
<th>Code-word</th>
<th>Segment</th>
<th>Code-word</th>
<th>Segment</th>
<th>Code-word</th>
<th>Segment</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$A_0$</td>
<td>21</td>
<td>00000-</td>
<td>42</td>
<td>+</td>
</tr>
<tr>
<td>1</td>
<td>$A_1$</td>
<td>22</td>
<td>0000+</td>
<td>43</td>
<td>+0</td>
</tr>
<tr>
<td>2</td>
<td>$A_{-1}$</td>
<td>23</td>
<td>0000-</td>
<td>44</td>
<td>+00</td>
</tr>
<tr>
<td>3</td>
<td>$A_2$</td>
<td>24</td>
<td>000+</td>
<td>45</td>
<td>+000</td>
</tr>
<tr>
<td>4</td>
<td>$A_{-2}$</td>
<td>25</td>
<td>000-</td>
<td>46</td>
<td>+00+</td>
</tr>
<tr>
<td>5</td>
<td>$A_3$</td>
<td>26</td>
<td>0+</td>
<td>47</td>
<td>+00-</td>
</tr>
<tr>
<td>6</td>
<td>$A_{-3}$</td>
<td>27</td>
<td>0-</td>
<td>48</td>
<td>+0+</td>
</tr>
<tr>
<td>7</td>
<td>$A_4$</td>
<td>28</td>
<td>0</td>
<td>49</td>
<td>+0-</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>29</td>
<td>0+0</td>
<td>50</td>
<td>++</td>
</tr>
<tr>
<td>9</td>
<td>00</td>
<td>30</td>
<td>0+00</td>
<td>51</td>
<td>+--</td>
</tr>
<tr>
<td>10</td>
<td>000</td>
<td>31</td>
<td>0+0+</td>
<td>52</td>
<td>--</td>
</tr>
<tr>
<td>11</td>
<td>0000</td>
<td>32</td>
<td>0+0-</td>
<td>53</td>
<td>-0</td>
</tr>
<tr>
<td>12</td>
<td>00000</td>
<td>33</td>
<td>0+</td>
<td>54</td>
<td>-00</td>
</tr>
<tr>
<td>13</td>
<td>000000</td>
<td>34</td>
<td>0-</td>
<td>55</td>
<td>-000</td>
</tr>
<tr>
<td>14</td>
<td>0000000</td>
<td>35</td>
<td>0-</td>
<td>56</td>
<td>-00+</td>
</tr>
<tr>
<td>15</td>
<td>00000000</td>
<td>36</td>
<td>0</td>
<td>57</td>
<td>-00-</td>
</tr>
<tr>
<td>16</td>
<td>0000000+</td>
<td>37</td>
<td>0-00</td>
<td>58</td>
<td>-0+</td>
</tr>
<tr>
<td>17</td>
<td>0000000-</td>
<td>38</td>
<td>0-0+</td>
<td>59</td>
<td>-0-</td>
</tr>
<tr>
<td>18</td>
<td>000000+</td>
<td>39</td>
<td>0-0-</td>
<td>60</td>
<td>--</td>
</tr>
<tr>
<td>19</td>
<td>000000-</td>
<td>40</td>
<td>0-+</td>
<td>61</td>
<td>--</td>
</tr>
<tr>
<td>20</td>
<td>00000+</td>
<td>41</td>
<td>0--</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

number of successive relative vectors. In poor transmission channels suffering from bit errors, choosing higher protection levels will lead to better picture quality, but obviously also to lower efficiency, owing to the insertion of redundant data. This approach can also be applied with MTC, without any restrictions.
6.4 Summary and Conclusions

Telewriting is a useful facility in tele-education systems. In this chapter a number of practical improvements of present telewriting systems were proposed concerning a colour fill facility, an efficient encoding algorithm and a method which supports editing of existing pages.

It has been shown that Differential Chain Coding, a technique used to encode telewriting, can be enhanced with a fill facility in such a way that graphical editing functions remain feasible. The additional coding information consists of only 3 elements, to be added to the DCC trace header. Thus, the ASCII-compatible transmission of line graphics proposed in Arnbak, Bons & Vieveen [89] can be readily extended to support coloured maps and other representations of areas with colour(s).

The extension or editing of on-screen-displayed objects by means of an electronic writing tablet implies an eye-to-hand coordination problem. A Relative Coordinate System (RCS) facility has been proposed by which the user is able to position handwritten graphics on screen at a precisely defined location.

DCC was shown to be a particular, but sub-optimal form of Modified Tunstall Coding. It is possible to optimize DCC by applying an extended set of codewords, which represent segments of relative or absolute vectors. It has been shown that a significant reduction of the coding rate results. This efficient method preserves DCC's typical robustness against transmission errors.
References in Chapter 6


Huffman, D.A. [52], "A method for the construction of minimum redundancy codes", *Proc. IRE*, Vol.40, pp.1098-1101


Kegel, A. & Bons, J.H. [77], "On the digital processing and transmission of handwriting and sketching", *Proc. EUROCON'77*, Venice, 1977, pp. 2.4.6.1-6


Lalor, G.C. [84], "The University of the West-Indies distance teaching project (UWIDITE)", *Higher Education* (L.A. and Caribb.), No.15, 1984


7 REDUCING END-TO-END DELAYS IN REAL-TIME MULTIMEDIA TELE-EDUCATION APPLICATIONS

7.1 Introduction

Real-time tele-education by means of multimedia technology enables universities to extend their educational facilities to remote students. Using standard PC networks and equipment to implement multimedia terminals offers the possibility to produce and control educational materials in a cost-effective way. The teacher usually supports his lecture by a number of 'screens' or 'pages' prepared in advance. These may indeed resemble compound pages, i.e., consisting of various data types like graphical and text objects, pictures, scanned overheads, animations and even short audio or video-tracks. During the session these pages are presented in a way similar to a conventional presentation of overhead sheets or slides as in use with face-to-face education. Moreover, the teacher is able to add or correct the presented items on-line. Beside the processing and transmission of these types of information, a tele-education system should provide two-way audio communication. In certain situations even real-time video communication (e.g., low frame-rate video) is a compelling requirement.

The production and display of these various types of information is not a major problem any more. The current generation of PCs and workstations offer enough facilities and performance for an efficient implementation of these functionalities. However, the main problem in many situations is still the available telecommunication bandwidth, which is constrained (e.g., at most 9600 bit/s in many operational circumstances) due to the analog public switched telephone network. Consequently, the transfer times of pages
resembling (compressed) pictures and other sub-Megabit objects, are not acceptable for real-time teaching. Table 7.1 lists typical transfer times for various pages at 2400, 9600 and 56000 bit/s. The listed pages contain only single data-type components, but in reality compound pages are frequently used. Although appropriate compression methods are assumed and image resolutions are moderate (but acceptable for tele-education purposes), it is clear that narrow bandwidths result in unacceptable delays.

Section 7.2 gives a survey of delay reduction methods which are applicable for the transmission of multimedia documents. These general methods suit a wide range of applications. However in the case of real-time tele-education more effective reduction is possible by alternative strategies, matched to the particular class of multimedia documents for this application. Section 7.3 describes this concept for point-to-point communication and multipoint systems. Section 7.4 addresses multicast configurations, characterized by large numbers of user stations serviced by one common tele-education channel.  

Table 7.1 Typical page transmission times for different data types as a function of transmission bitrate (2400, 9600 and 56000 bit/s)

<table>
<thead>
<tr>
<th>Data type</th>
<th>Compression</th>
<th>Page size kBy</th>
<th>Transmission time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2400</td>
<td>9600</td>
</tr>
<tr>
<td>Text</td>
<td>-</td>
<td>5</td>
<td>16.7</td>
</tr>
<tr>
<td>Vectorised graphical</td>
<td>-</td>
<td>5</td>
<td>16.7</td>
</tr>
<tr>
<td>Handwritten graphics</td>
<td>DCC</td>
<td>10</td>
<td>33.3</td>
</tr>
<tr>
<td>Raster scanned 2-lev.</td>
<td>RLC</td>
<td>8</td>
<td>26.7</td>
</tr>
<tr>
<td>Colour picture</td>
<td>JPEG</td>
<td>50</td>
<td>168</td>
</tr>
<tr>
<td>5 s Telephone quality</td>
<td>PCM</td>
<td>40</td>
<td>133</td>
</tr>
<tr>
<td>5 s Video track</td>
<td>MPEG</td>
<td>875</td>
<td>2917</td>
</tr>
</tbody>
</table>

[Compression methods:
DCC: Differential Chain Coding (Kegel & Bons [77])
RLC: Runlength Coding (CCITT Group IV facsimile)
JPEG: Image coding standard (Wallace [91])
PCM: Pulse code modulation (Jayant [74])
MPEG: Moving Pictures coding standard (Le Gall [91])]

---

1 This work has been presented at the SPIE/EOS International Symposium on Fiber Optic Networks and Video Communications, Berlin, April 1993 (Algra [93])
7.2 End-to-end delay reduction methods

Rosenberg et al. [92] identify three methods to shorten page transfer times: media compression, structural transmission and time shifting. These methods do not impose constraints on the authoring process and in general allow any object type, object and document size, and document structure. However, here we will distinguish two additional methods, which both influence the way of authoring: document structuring rules and object type selection (refer to Table 7.2):

1) Document structuring rules
   Nodes\(^2\) in the document which contain one or more objects with a large data size and consequently a long transfer time, should only be reached via nodes of which the information content requires some time for interpretation by the user. The system uses this period to transfer the amount of data corresponding to the node with the large object(s). Hardt-Kornacki & Ness [91] describe the example of a video sequence which should be preceded by a text node which informs the user on the subjects presented by that sequence.

2) Object type selection
   This method implies that the authoring process is accomplished by accepting a certain restricted subset of possible object types and sizes, in order to reduce transfer times.

3) Media compression
   Medium compression uses medium-specific and cognitive properties to compress bandwidth intensive media like still images and full-motion video. JPEG, MPEG and H.261 are examples of medium compression techniques (Wallace [91], Le Gall [91], Liou [91]).

4) Structural transmission
   Structural transmission saves bandwidth by exploiting the document's structure. Two kinds of document structure are available: semantic and media structure.

\(^2\) A node in a multimedia document is defined as the smallest entity of information which can be selected for retrieval
Table 7.2 Survey of end-to-end delay reduction methods in multimedia communications

<table>
<thead>
<tr>
<th>Authoring rules</th>
<th>Data processing &amp; handling</th>
</tr>
</thead>
<tbody>
<tr>
<td>o Document structuring</td>
<td>o Media compression</td>
</tr>
<tr>
<td>o Object type selection</td>
<td>o Structural transmission</td>
</tr>
<tr>
<td></td>
<td>- Semantic structure</td>
</tr>
<tr>
<td></td>
<td>- Media structure</td>
</tr>
<tr>
<td></td>
<td>o Time shifting</td>
</tr>
<tr>
<td></td>
<td>- Downloading</td>
</tr>
<tr>
<td></td>
<td>- Database synchronisation</td>
</tr>
<tr>
<td></td>
<td>- Delivery control</td>
</tr>
<tr>
<td></td>
<td>- Pretransfer</td>
</tr>
</tbody>
</table>

4a) **Semantic structure** uses author-specified semantics to eliminate redundant transmission. For example, consider a presentation with a fixed pixmap background upon which other media appear and disappear. The background is sent only once. The display machine understands the semantics of a background and will repaint parts of the pixmap when necessary, without the need for communication with the document sender.

4b) **Media structure** reduces the amount of information transmitted by exploiting the structure of specific media. For example, consider a text block in a compound document, consisting of a rectangular coloured area and containing 50 characters in a 10-point Times Roman font. If this were sent as a pixmap, it might require tens of kBytes. On the other hand the block could be sent as structural information: a description of the rectangle, its colour, the character font, and the characters themselves. This representation would require less than 100 bytes (Rosenbergh *et al.* [92]).

5) **Time shifting**
Time shifting means that the page data are delivered to the target system before they are required for presentation. We will distinguish various types of time...
7.2 End-to-end delay reduction methods

shifting: downloading, database synchronisation and delivery control.

5a) Downloading, a well-known approach to time shifting, involves the transmission of all the pages separately by a file transfer service, prior to the actual lecture.

5b) Database synchronisation: In Bruyne [91], and Novara, Collet & Vaernes [91] a tele-instruction system is described where two identical multimedia databases are kept synchronized during the session, by the use of a data link. This method implies that a user at one location is able to generate retrieval commands which effect both data-bases in the same way. Furthermore, the described system facilitates pointing and writing (by means of an electronic writing tablet with a handheld stylus) on a shared screen. A second communication circuit is used for two-way audio. A similar system is reported by Rosenthal [87], who describes the distance teaching system of the University of the Virgin Islands. In this case the operating system of a remote microcomputer is synchronised with an identical microcomputer at the disposal of the lecturer. The approach requires that the same information (software and data) is permanently stored at both sides. A two-way audio circuit completes the system.

5c) Delivery control: a delivery control facility (Hardt-Kornacki & Ness [91], Loeb [92]) is a network advice service, which contributes towards minimizing network latency, maximizing network utilisation and optimizing the overall quality of service which is provided to the end-user. The adaptive allocation of network resources and the transmission in advance of objects with a high probability of future access are the means of control, based upon information like end-user navigation behaviour, document structure, network- and display capacity.

In this chapter we will discuss time shifting techniques matched to real-time tele-education applications. It will be shown that it is necessary to develop an alternative approach to achieve more effective latency reduction with less impact on organisational and cost aspects.

Downloading is a straightforward approach, but suffers a number of drawbacks:

a) The increased link connection time per session
b) Extra attention and organisation are necessary, both at the transmit and the reception sites
c) Depending on the quality of the organisational and technical infrastructure, there is a certain risk that the file transfer may not be completed correctly, which may result in postponement of the lecture
d) In the case of multipoint systems the frequency of occurrence of the latter problem increases with the number of reception sites.

Database synchronisation requires well structured and operationally available educational information in the various databases at all sites, before the session starts. In other words, a teacher is not able to make 'last minute' corrections and additions to his lessons, and the information should be frozen a certain time before use. Such a reduction of flexibility in lecture preparation deviates significantly from what is common practice among university teachers. Moreover, database synchronisation requires considerable multi-site configuration control. Hence this method has undesirable consequences for staff resources, way of teaching and financial budgets. The described delivery control method was proposed (Hardt-Kornacki & Ness [91]) for interactive documentation delivery, originating from an information provider (e.g., a database), to an end-user via a network facilitating dynamic capacity allocation. In the case of real-time tele-education, network capacity often is not variable (fixed bitrate). Furthermore the specific document access graph of a tele-education presentation differs significantly from the more general type as supposed in the case of a delivery control facility (refer to 3.2.2.2). Nevertheless, some of the concepts of delivery control are applicable to tele-education.

The remaining part of this chapter deals with a fourth approach to time shifting which we refer to as pretransfer. The method employs periods of low channel utilisation during the session for the automatic transmission of the page sequence in background. This process is not observed by the user, since the transmission of objects related to the real-time tele-education process will have absolute priority. Consequently, pretransfer application protocols result in virtual zero transfer times (VZT) and more efficient use of transmission resources.

The presented solution is not only suitable for typical tele-education systems, but could also enhance other tele-instruction applications with limited link connection time. One example is the envisaged interactive communication between astronauts in future space laboratories, such as COLUMBUS, and scientists on their user homebase (UHB) (Pronk et al. [92]). The on-board crew will be responsible for the execution of numerous scientific experiments originating from hundreds of scientists. In the event of unexpected problems related to a particular experiment or equipment on-board, and the consequential need for instruction of the crew by the investigator, a multimedia communication tool with a pretransfer function will be very effective. Both space-link connection time and crew time are scarce and thus expensive resources, demanding well prepared communication sessions. Pretransfer results in zero waiting times with respect to uplinked pages, even in the case of a narrow bandwidth link to the UHB.
7.3 Point-To-Point Pretransfer Protocol

Assume two stations interconnected by a point-to-point link. Although either of the stations could be the originator of the call, always one of the two is permanently designated as the primary (station) and the other one as the secondary one. In the tele-education situation, it is the teacher who has the primary station at his disposal. When necessary, the primary station is able to overrule almost all the actions of the secondary one, if initiated during the session. For example, the former determines if and when the latter is allowed to modify the screen information. The pretransfer information stream is limited to one direction: from primary to secondary. In the following considerations only the application layer of the communication software is involved. During call set-up the (lower) session layer already has initiated and established a reliable connection, which serves the application layer with a constant capacity channel 3.

In this section a point-to-point pretransfer protocol is developed which automatically sends files to the secondary station, without user involvement. The user does not even notice this process, since there is no visible impact on system performance. The files sent are the pages of the current presentation, which defines a multimedia document consisting of $N$ pages, sequentially numbered from 0 to $N-1$. Note that the teacher is not forced to present the pages according to this sequential order, but that any arbitrary order may be followed.

7.3.1 Requirements

The protocol should fulfil (or support) the following requirements:

a) Obviously, the pretransfer process (PRP) should start as soon as the connection between the stations is established.

b) The process is not noticeable by the users (unless they explicitly request status information). This requirement includes unobservable system performance degradation.

c) The secondary station should be protected from disk overload. Insufficient storage

---

3 For the division of communication tasks into layers according to the OSI-model (Open Systems Interconnection) and the associated terminology, the reader is referred to the literature (e.g., MacKinnon, McCrum & Sheppard [90])
capacity results in a message to the user of the primary station, which includes the current status. Then the PRP stops sending files.

d) After an accidental link dropout, the PRP should resume at the point where it arrived before the dropout.

e) If the primary station retrieves a page for display at the screen, the PRP should check its status information and decide if the complete page should be sent or only a copy of the retrieval command. In the former case, the primary user should be notified on the completion of the page transmission process.

f) The PRP should only process files which belong to the current presentation. In order to facilitate correction of a wrong presentation choice, the PRP prompts the user to immediate confirm switching of pretransfer to another presentation as soon as selected. This avoids unnecessary file transmission if the teacher wants to switch to another presentation temporarily (e.g., to search for a picture, a graph, etc.).

g) The process can be reset by the primary user, which implies that already transferred files will be erased.

h) At any time, the user may request status information from the PRP, consisting of presentation name, number of transferred pages, total number of pages and the amount of free memory left for pretransfer at the secondary's storage medium.

7.3.2 The Protocol

The tele-education application contains two communication-related processes. First, the high-priority real-time communication of objects, pages and commands, which immediately effect the displayed information (RTC). Second, the low-priority pretransfer communication (PTC). Both processes are controlled by the User Interface (UI) and associated functions, and make use of the services of lower layers for the execution of communication tasks (Fig. 7.1). The communication part of the PTC, the Application Entity (AE), should be distinguished from other functions of the PTC like file-handling and information exchange with the UI. Hence there is no explicit boundary between the Application Layer User (ALU) and the Application Entity; the separation is merely conceptual. Further, the PTC can be split-up in two sub-processes: a lower-level process for the automatic execution of the protocol, including file-handling and transfer (Pretransfer Machine, PTM) and a higher level process (Pretransfer Supervisor, PTS). The latter interprets commands from the UI and takes the corresponding actions to control the PTM. Dividing the PTC in two sub-processes results in a well-structured design, and a clear interface between the two (Fig. 7.2). The application protocol between the P-PTM and the S-PTM only requires four primitives (Table 7.3). The operation of the protocol
7.3 Point-to-point pretransfer protocol

Fig. 7.1 Communication architecture

Fig. 7.2 Pretransfer application protocol (point-to-point). The order of the protocol messages is indicated by the number between brackets.

follows from Fig. 7.2 and the state transition diagram of Fig. 7.3. To start pretransfer, the P-PTS issues an initiate(id) message to the P-PTM. The parameter id is the identification of the presentation which should be transferred. Then the P-PTM requests information from the S-PTM on available storage capacity and on already transferred parts of the presentation, by a request(id) message. The S-PTM may respond with an accept(id,k) or a reject(id,k,s) message. The parameter k denotes the number of pages
which are already present at the secondary station. The parameter $s$ informs on the reason for the rejection (memory full, overflow, or a pretransfer disable command by the secondary user). Upon receipt of an accept message, the P-PTM starts transmitting the pages of the presentation, numbered $k$, $k+1$, ..., $N-1$. As long as the S-PTM is able to store the pages, no feedback messages are transmitted. However, as soon as a page cannot be stored, the S-PTM will send a reject message, and the P-PTM stops transmission. A reject message is notified to the P-PTM and will result in a message to user.

Fig. 7.4 depicts the STD of the P-PTS. It is easily verified that this structure fulfils the requirements given in 7.3.1. The S-PTS is a passive interface between the UI and the S-PTM and is used only for transferring status information. After a link drop-out, the
described protocol will always start with a status information exchange, which avoids unnecessary transmission of information already present at the secondary station. Furthermore the primary user is able to force a reset. This will result in erasure of all transferred files at the secondary station, initiated by a reset protocol message. During the time of the session the primary user will select and retrieve the various pages of the presentation for display, but not necessarily in sequential order. Every time such a retrieval command is given, the higher level part of the application (for convenience we will simply speak of the UI) requests the P-PTS if the concerning page data should be
transmitted completely in real-time or if pretransfer already was completed. In the latter case only a retrieval command should be transmitted. Therefore a call to a P-PTS service is executed, with the page number as a parameter. Let \( k \) denote the number of already transferred pages, and \( P(j) \) the page with number \( j \). Then the function can be defined as

\[
\text{if } j \geq k \text{ then } \text{TRANSMIT } P(j) 
\]  

(7.1)

### 7.3.3 Multipoint Pretransfer Protocols

The presented method is also applicable to multipoint configurations. The primary station only needs to start a separate pretransfer process for each secondary. These processes negotiate and handle pretransfer with all the secondary stations individually. Hence, multipoint pretransfer is a straightforward extension of the point-to-point case. However, one complicating issue should be discussed. The decision process which must be carried out each time when a new page is to be displayed, should be based on status information of all connected secondary stations. If one or more stations did not yet receive the page data by pretransfer (this may be due to bandwidth or performance constraints, or a late establishment of the connection), the data has to be transmitted in real-time, and the teacher has to take into account the resulting page transfer time. Let there be \( M \) secondary stations, indexed by \( i, i \in \{1, \ldots, M\} \). Let \( k_i \) denote the number of transferred pages for station \( i \). Then the decision whether to transmit page \( P(j) \) in real-time, has to be taken according to

\[
\text{if } j \geq \text{MIN } \{ k_i, i \in \{1, \ldots, M\} \} \text{ then TRANSMIT } P(j) 
\]  

(7.2)

### 7.4 Multicast Pretransfer

In multicast situations there is no return link from the secondary station to the primary station. In tele-education applications only a phone-in service may be implemented to allow voice feedback. Thus all pages should be transmitted continuously, by an endless loop, via the pretransfer channel (as in a teletext protocol, IEEE [79]). This approach guarantees:
a) that a secondary station which for some reason has not realized a connection at the session start, but only later on, it will still collect all pages of the session;
b) that such a 'late' station will also, after a certain time, reach the VZT state.

Let a session consist of \( N \) consecutive pages \( P_n, n \in \{0,1,2,...,N-1\} \). These pages are transmitted via the pretransfer channel according to a certain schedule. The pretransfer channel rate is \( B \) bit/s. The data size of page \( P_n \) is denoted as \( L_n \) (bits). Hence the transmission time, \( T_s(n) \), of page \( P_n \) through the pretransfer channel equals \( L_n/B \). Assume that the primary station changes between the \( N \) consecutive pages of the session at instants \( t_n, \ t_n \in \{t_0, t_1, \ldots, t_N\} \), and that the session is closed at \( t_N \). [For simplicity reasons, we shall ignore the possibility of a lecturer being able to deviate from the predefined order of the pages] Let the pages be displayed at the secondary terminal at instants \( t_n^*, \ t_n^* \in \{t_0^*, \ t_1^*, \ldots\} \). The transfer time of page \( P_n \) is defined as \( T_\tau(n) = t_n^* - t_n \). Assume that a secondary station connects to the primary multicast data signal at an arbitrary instant \( t' \), with \( t_{n-1} < t' < t_n \). Let page \( P_k \) be the first page which is displayed at the secondary terminal, say at instant \( t_k^* \). Thus \( k \geq n-1 \). Then \( T_\delta = t_k^* - t' \) defines the time that the user at this secondary terminal has to wait before the first page is displayed. We will refer to \( T_\delta \) as the log-in time. After a certain time, the secondary terminal will have collected all pages yet to be displayed, via the pretransfer mechanism. Eventually, this results in virtual zero transfer times. Let page \( P_m \) be the first page which is displayed at the secondary terminal with a \( T_\delta(m)=0 \). \( T_\sigma = t_m^* - t' \) defines the time between connection establishment and the instant at which virtual zero transfer times are achieved, and will be referred to as the synchronisation time. The optimal pretransfer schedule should result in minimal values of \( T_\delta \) and \( T_\sigma \).

Four possible pretransfer schedules will be analyzed, comparing worst case values of \( T_\sigma \) and \( T_\delta \), as a function of \( B \), under the following two assumptions:

1) \( L_n = L, \ n \in \{0,...,N-1\} \) 
   Hence \( T_s(n) = L/B = T_s \)

2) \( t_n - t_{n-1} = t_n/N = T_m, \ n \in \{1,2,...,N-1\} \)

The four schedules to be analyzed, are defined as:

A: Non Pre-emptive Scheduling (NPS)

The \( N \) pages are transmitted through the pretransfer channel according to a round-robin type cycle: \( 0,1,\ldots,N-1,0,1,2\ldots,N-1,0,1,2\ldots \) etc.
B: **Current Page Pre-emptive Scheduling (CPPS)**
This schedule equals the NPS scheduling type, but with the exception that the round-robin cycle is temporarily suspended at each instant $t_n$ for the transmission of page $P_n$. When transmission of the latter is completed, the cycle is resumed. In other words, two pretransfer processes exist: a background process consisting of the repeated $N$-pages cycle and a higher-priority foreground process which involves the transmission of the current page.

C: **Next Page Pre-emptive Scheduling (NPPS)**
This schedule is similar to CPPS, except that the foreground process starts to transmit page $P_{n+1}$ (instead of $P_n$) at instant $t_n$.

D: **Current & Next Page Pre-emptive Scheduling (CNPPS)**
This schedule is similar to the previous one, except that the foreground process starts to transmit both page $P_n$ and page $P_{n+1}$ at instant $t_n$.

Fig. 7.6 visualizes these four scheduling types. For convenience, we will use time values normalized on $T_u$ and introduce $\beta = T_e/T_u = L/(BT_u)$. The worst-case time-interval in which all $N$ pages are transferred by the background process, is referred to as $T_b$.

---

**Fig. 7.6 Possible pretransfer scheduling methods for multicast situations (N=5)**
7.4 Multicast pretransfer

A: Non Preemptive Scheduling (NPS)

At instant $t' + N\beta$ all pages are received by the secondary terminal. Hence the first page is displayed at $t_k^*$, with $t' + \beta \leq t_k^* \leq t' + N\beta$. Then the log-in time becomes:

$$T_b = N\beta$$  \hspace{1cm} (7.3)

Let $t_k < t' + N\beta \leq t_{k+1}$. Then $T_r(k) = t_k^* - t_k > 0$ is worst case, but $T_r(k+1) = t_{k+1}^* - t_{k+1} = 0$. Or: $t_{k+1}^* - 1 < t' + N\beta$. Thus $t_{k+1}^* - t' < 1 + N\beta$.

Hence

$$T_a = 1 + N\beta$$  \hspace{1cm} (7.4)

Obviously,

$$T_b = N\beta$$  \hspace{1cm} (7.5)

B: Current Page Pre-emptive Scheduling (CPPS)

The foreground process results in an upper limit $1+\beta$ for $T_b$. For $N\beta < 1-\beta$ the instant $t_k^*$ is delayed $\beta$ (worst case) as compared to NPS-scheduling.

$$T_b = \begin{cases} 
(N+1)\beta & \text{if } N\beta < 1 - \beta \\
1 + \beta & \text{otherwise} 
\end{cases}$$  \hspace{1cm} (7.6)

For $T_a$ the derivation of NPS-scheduling applies, except for a slight extension $\lfloor N\beta/(1-\beta) \rfloor$ due to the foreground process.

$$T_a = 1 + N\beta + \lfloor \frac{N\beta}{1-\beta} \rfloor$$  \hspace{1cm} (7.7)

It is easily verified that

$$T_b = (N+1)\beta$$  \hspace{1cm} (7.8)

C: Next Page Pre-emptive Scheduling (NPPS)

In this case the foreground process results in an upper limit for $T_b$, which is equal to 2. For $N\beta < 1 - \beta$ the situation is similar to CPPS-scheduling. For $1 - \beta \leq N\beta < 2(1-\beta)$ the
instant \( t_k^* \) is delayed \( 2\beta \) (worst case) as compared to NPS-scheduling:

\[
T_b = \begin{cases} 
(N+1)\beta & \text{if } N\beta < 1-\beta \\
(N+2)\beta & \text{if } 1-\beta \leq N\beta < 2(1-\beta) \\
2 & \text{otherwise}
\end{cases} 
\]  
\( \quad (7.9) \)

For \( T_o \) the same situation applies as with NPS and CPPS-scheduling, except that an upper limit exists: Let \( t_k < t' \leq t_{k+1} \). Page \( P_{k+2} \) is received at \( t_{k+1} + \beta \).

Hence \( T_o(k+2) = t_{k+2}^* - t_{k+2} = 0 \). Thus \( t_{k+2}^* - t' = t_{k+2} - t' < t_{k+2} - t_k = 2 \).

\( T_o \) becomes:

\[
T_o = \begin{cases} 
1 + N\beta & \text{if } N\beta < 1-\beta \\
2 & \text{otherwise}
\end{cases} 
\]  
\( \quad (7.10) \)

\( T_b \) is equal to CPPS

\[
T_b = (N+1)\beta 
\]  
\( \quad (7.11) \)

**D: Current & Next Page Pre-emptive Scheduling (CNPPS)**

It is easily verified that the upper limit for \( T_b \) is equal to 1. For \( N\beta < 1-2\beta \) the instant \( t_k^* \) is delayed \( 2\beta \), worst case, as compared to NPS-scheduling:

\[
T_b = \begin{cases} 
(N+2)\beta & \text{if } N\beta < 1-2\beta \\
1 & \text{otherwise}
\end{cases} 
\]  
\( \quad (7.12) \)

For \( T_o \) the same situation applies as for NPPS-scheduling, except that the upper limit is changed to a lower value: if \( t' \) lies within the interval \( (t_k, t_k+\beta) \) then page \( P_{k+1} \) is received at \( t_k + 2\beta \). Hence \( t_{k+1} = t_{k+1}^* \) and thus \( T_o \leq 1 \). If \( t' \) lies in the interval \( (t_k+\beta, t_{k+1}) \) then, worst case, page \( P_{k+1} \) is not received in time and consequently \( t_{k+1}^* > t_{k+1} \). In other words then \( T_o = 2-\beta \). Hence

\[
T_o = \begin{cases} 
1 + N\beta & \text{if } N\beta < 1-2\beta \\
2-\beta & \text{otherwise}
\end{cases} 
\]  
\( \quad (7.13) \)
In this case $T_b$ becomes

$$T_b = (N+2)\beta$$  \hspace{1cm} (7.14)

In Fig. 7.7 all these results are plotted as a function of $\beta$, where it is assumed that $N \gg 1$. The best results are obtained with CNPPS-scheduling.

![Fig. 7.7 Log-in time $T_5$ and synchronization time $T_\sigma$ as a function of $\beta$ for the four different scheduling methods](image)

CNPPS-scheduling can be generalized in the following way. Assume that the foreground-process is resumed at instants $t_k + r\beta$, $t_k + 2r\beta$, $t_k + 3r\beta, \ldots$, $t_k + (\lceil 1/(r\beta) \rceil - 1)r\beta$, under the condition that $r > 2$. This form of CNPPS-scheduling will be referred to as CNPPS-II. The longest interval without active foreground process occurs just before instant $t_{k+1}$ and its length equals $1 - (\lceil 1/(r\beta) \rceil - 1)r\beta - 2\beta$. Assume that $r \ll N$. For $\lceil 1/r\beta \rceil - 1 \geq 1$, or: $\beta \leq 1/2r$, CNPPS-II applies and $T_5$ changes as compared to CNPPS-I. When $\beta > 1/2r$ then $N\beta > 1-2\beta$, since $r \ll N$. Hence
\[ T_\delta = \begin{cases} 
1 - ((1/r\beta - 1)\beta & \text{if } \beta \leq \frac{1}{2r} \\
1 & \text{otherwise}
\end{cases} \] (7.15)

If \( r' > t_k + (\lfloor 1/(r\beta) \rfloor - 1)\beta + \beta \) then \( t_{k+1}^* > t_{k+1} \).

Hence \( T_\sigma = 1 + 1 - (\lfloor 1/(r\beta) \rfloor - 1)\beta \beta = 2 - (\lfloor 1/(r\beta) \rfloor - 1)\beta - \beta \)

\[ T_\sigma = \begin{cases} 
2 - ((1/r\beta - 1)\beta - \beta & \text{if } \beta \leq \frac{1}{2r} \\
2 - \beta & \text{otherwise}
\end{cases} \] (7.16)

The choice of \( r \) influences the length of the background period \( T_b \). On an interval \((t_k, t_{k+1}]\) the foreground is active during \( \lfloor 1/r\beta \rfloor 2\beta \) units of time. The remaining time is left for the background process, which then transmits \((1 - \lfloor 1/r\beta \rfloor 2\beta)\beta \) pages. Hence \( T_b = \beta N/(1 - \lfloor 1/r\beta \rfloor 2\beta) \). For \( T_b \) an upper limit can be derived: \( T_b \leq \beta N(r/2r) \). In Fig. 7.8 \( T_\delta \) and \( T_\sigma \) are plotted for \( r = 3 \). For a typical case with a session of 25 pages and duration of 50 minutes, with \( L = 10 \) kbit, \( T_\sigma \) and \( T_\delta \) are plotted as a function of \( \beta \) for CNPPS-II scheduling in Fig. 7.9.

It can be concluded that CNPPS-II is an effective approach to pretransfer in the case of multicasting of tele-education sessions.

![Diagram](image)

Fig. 7.8 Log-in time \( T_\delta \) and synchronization time \( T_\sigma \) as a function of \( \beta \) for CNPPS-II
7.4 Multicast pretransfer

![Graph showing $T_\delta$, $T_\sigma$, and $T_\alpha$](image)

Fig. 7.9 Log-in time $T_\delta$ and synchronization time $T_\sigma$ for CNPPS-I and CNPPS-II as a function of pretransfer channel rate ($N=25, L=10^3, T_\alpha=120$ s).

7.5 Summary and conclusions

Real-time tele-education by means of multimedia technology enables universities to extend their educational facilities to students on remote locations. In many situations, transmission of multimedia documents is still limited due to network bandwidth constraints (e.g., the analog telephone network). To reduce end-to-end delays, various approaches may be adopted: media compression, structural transmission and time shifting.

This chapter proposed a time shifting method, referred to as pretransfer, which exploits periods of low channel utilisation during the session for automatic transmission of page sequences in background. These compound pages were assumed to be prepared before the start of a session. When the lecture is in progress, arbitrary objects and pages can be added on-line. The pretransfer process is not noticeable by the user and results in virtual zero transfer times. Application protocols have been discussed for both point-to-point and multi-point configurations.

In multicast situations, which are characterised by a large (unnumbered) number of reception stations, pretransfer can be realised by means of appropriate page-scheduling
methods. The time from the moment a station links in to the multicast signal to the establishment of the virtual zero transfer time mode for that station, is referred to as the synchronisation time $T_o$. Various scheduling methods were analyzed, and compared with respect to $T_o$ (worst-case values) as a function of channel rate, number of pages and page size. It has been concluded that pretransfer methods are an effective means to further reduce end-to-end delays in typical tele-education applications.
References in Chapter 7

Algra, T. [93], "End-to-end delay reduction in narrow bandwidth real-time multimedia tele-
education applications", Proc. EOS/SPIE Int. Symp. on Fiber Optic Networks and Video Com-

Bruyne, H.C.D. de [91], "Multimedia/Teletraining Experiment", Study performed by CAT
Benelux for ESA, ESA Study contract report, ESA-CR(P)-3315, August 1991

Hardt-Kornacki, S.L., L.A.Ness [91], "Optimization model for the delivery of interactive multime-


Le Gall, D.J. [91], "MPEG: a video compression standard for multimedia applications", Comm.
of the ACM, Vol.34, No.4, pp. 46-58, 1991

Jayant, N.S. [74], "Digital coding of speech waveforms: PCM, DPCM, and DM quantizers,

Kegel, A., J.H. Bons [77], "On the digital processing and transmission of handwriting and sket-
ching", Proc. EUROCON'77, Venice, pp.2.4.6.1-6, 1977

Liou, M.L. [91], "Overview of the px74 kbit/s video coding standard", Comm. of the ACM,
Vol.34, No.4, pp. 59-73, 1991

Loeb, S.L. [92], "Delivering interactive multimedia documents over networks", IEEE Comm.
Mag., Vol. 30, No.2, pp. 52-59, May 1992

MacKinnon, D., W. McCrum & D. Sheppard [90], An Introduction to Open Systems Intercon-

Novara, M., J.Collet, R.J.Vaernes [91], "The ISEMSI experiment: Europe's first simulation campa-

Prong, C.N.A., N.Koopman, D.de Hoop [92], "Development concept for Dutch User Support",
Proc. 43rd Congress of the International Astronautical Federation, Washington, Sept. 1992

Rosenberg, J., R.E.Kraut, L.Gomez, C.A.Buzzard [92], "Multimedia communications for users",

Rosenthal, L. [87], "The Caribbean Center Telecommunications Project", University of the
Virgin Islands, February 1987

Wallace, G.K. [91], "The JPEG still picture compression standard", Comm. of the ACM, Vol.34,
No.4, pp. 31-44, 1991
8 A PC-BASED REAL-TIME Multimedia Tele-Education System

8.1 Introduction

Modern telematic technology enables the extension of the educational facilities of universities to remote locations. Real-time tele-education based on multimedia concepts permits the realization of effective distance courses. Application in third-world countries usually implies a number of specific conditions and constraints. This concerns the available telecommunication infrastructure, the local level of serviceability regarding complex equipment, and the limitations of the educational institute involved. This chapter describes the development of a real-time tele-education system which can be used in situations where a requirement exists for PC-based implementation in combination with narrowband telecommunication links. The requirements stated in the next section are derived from analyzing the case of the University of the Netherlands Antilles (UNA) at Curaçao. Curaçao belongs to an archipelago of small islands with mutual distances up to 800 kilometres, located in the Caribbean Basin. Only Curaçao accommodates an institute for higher education. This particular case represents a broad range of similar situations, as has been pointed out in Chapter 3. Prior work related to real-time tele-education is addressed in Section 3 of this chapter, while the subsequent sections describe the system developed and the results of a classroom experiment performed to validate the assumptions and design criteria adopted in this thesis.
8.2 Objectives, requirements, and constraints

The main objective for a tele-education service may be stated as follows. The system should enable effective distance teaching by means of telecommunication technology. This should result in a significant higher level of participation in the educational programs of the university by inhabitants of the remote areas considered.

This objective leads to the formulation of requirements for a tele-education system. The key requirements are listed below. For a more detailed requirements analysis the reader is referred to Chapter 3.

1) Real-time interactive teaching
In tele-education the same methods as in the normal face-to-face teaching situation should be feasible in order to guarantee a sufficiently high level of learning effectiveness. The system should enable interactive communications between teacher and students, since feedback from student to teacher is of vital importance. However the long turn-around time in correspondence-based distance education should be avoided. This requires real-time interactivity with adequate system facilities to enable, and stimulate these interaction processes.

2) Media
Interactive communication patterns in education usually consist of three main components: audio communication for information transfer and feedback; educational materials to support this information transfer (blackboard, videotapes, printed materials, etc.); and visual communication. The latter enables non-verbal communication, including observation of the students' attitude. The tele-education system should support all these three components by adequate media. This implies a two-way audio and a two-way visual communication system, and further a medium which enables information transfer corresponding to blackboard, overhead-projector, slides, and other visual items as used in normal face-to-face teaching.

3) Suitability to different levels of user experience
It should be possible to apply the system in a low-level way for incidental users as a guest lecturer, without the need for extensive training. Still, for experienced users the full advantages of a well-balanced and effective computer application should be taken. This may include filing, editing and library facilities.
4) **Multipoint communications and floor control**

The system should be suitable to teach multiple remote groups simultaneously. The teacher should have control over the feedback signals coming from the remote groups in order to allow a structured and well-ordered interaction process. This mode of operation will be referred to as the primary/secondary mode. A password procedure should prevent unauthorized access to this mode. By default the system is in the peer-to-peer mode without imposing an artificial hierarchy on participants. The latter is more appropriate for teleconferencing.

5) **Structured presentations**

It should be possible both to organize and deliver highly structured presentations, and to introduce visual material out of sequence.

6) **Editing and filing**

The system should allow interactive updating and annotation of all visual media by all the participants ('shared screen'). Further, appropriate filing and retrieving functions should be added. This enables efficient off-line production and management of educational materials.

7) **Performance issues**

Real-time interactivity requires minimal information transfer times for all the media applied. Shared screens should be kept synchronized under all circumstances.

The above requirements are to be met subject to constraints. The envisaged situation of a small-scale university in a third-world country and a limited number of remote communities imposes the following constraints and limiting conditions.

1) **Educational and organizational aspects**

The system should allow to continue existing teaching methods. Additional human resources for development and operation is not feasible for a small-scale university.

Moreover, to keep the psychological threshold for system introduction low, this novel technology should not be combined with new working methods (Bates [88], Arnbak [88], Carey [89]). Further the system should permit small-scale implementation, thus reducing financial and educational risks, and allowing smooth transfer to a more extensive realization.

2) **Technical aspects**

In many situations only narrowband analog telecommunication links are possible. Hence it should be possible to operate the system at any bitrate \( \geq 1200 \text{ bit/s} \). Still, 9600 bit/s should be considered as a design driver, since this capacity can be achieved in many situations.
The available level of service and maintenance regarding complex equipment may be insufficient in many cases. Therefore the use of standard equipment such as PCs and associated peripherals is recommended. A maximum of functionality should be realized in software.

3) Budgetary aspects
A small-scale university (< 1000 students) and target groups of a limited size require low-cost solutions. This concerns both investment and operational costs of equipment, telecommunication links, and staff required (teachers, technicians, and management). With respect to equipment preferably existing microcomputers, peripherals and local networks should be used where possible. This leads to a system based on Personal Computers.

In the following sections we discuss previous work related to real-time tele-education, then describe a system to meet the above mentioned objective and requirements.

8.3 Related developments

Two fields of prior work are of particular importance to the objective and requirements stated above. The first will be referred to as conventional real-time tele-education; the second is visual teleconferencing.

Conventional real-time tele-education encompasses systems which do not apply computer technology. Audio-conferencing (Winders [85]) and ITV (DeSio [92]) have been used successfully since years. ITV-systems mostly employ one-way video with audio return. A disadvantage of ITV is the large bandwidth required. If this is not available, slow scan video (SSTV) may be applied, extended with a telewriting facility. For example the distance teaching network of the University of the West-Indies employs regular analog telephone circuits for audio, SSTV and telewriting transmissions (Lalor & Marrett [86]). A similar system is reported by McConnell [86]. An early example of audio-graphic tele-education is the 'scribaphone' (Bordewijk [78]). Here a digitally encoded handwriting signal is transmitted together with an audio signal over a single analog telephone circuit, using a special multiplexing device. Severe drawbacks of these audio-graphics systems, whether including a variant of SSTV or not, are: 1) the application of special-purpose hardware; 2) long picture transfer times due to narrow signal bandwidths; 3) lack of integration with a computer system to support efficient production and management of educational materials.
The advent of the Personal Computer in the Eighties paved the way to eliminate these disadvantages. Early PC-based real-time tele-education systems are described in Sciglano & Centini [85] and Rosenthal [87]. The former applies a simple text-based communication method among students and a teacher, linked together by PDN connections. The latter combines two-way audio with a communication scheme involving fully synchronized microcomputers. The remote computer is slaved to the teacher’s computer, enabling simultaneous execution of application programs at both sites. Gao, Li & Li [89] describe a PC-based system realizing simultaneous data and audio transmission over a single PSTN circuit by a special modem. The data signal may transfer stills (from a frame-grabber, connected to a video camera or a VCR) or computer-generated text and graphics. Either option can be selected by a manual switch. Normally, still pictures are down-loaded prior to the lecture, and may be called up when necessary. Hence the video stills cannot be transmitted simultaneously to the audio and/or the graphics signals. Another variant is reported by McDonnell [92]. Here, by a similar voice-data modem, telewriting as well as computer commands may be relayed. The latter enables simultaneous retrieval of PC-graphics at all remote locations, provided that these data have been downloaded in advance to the session.

Visual teleconferencing has traditionally been based on special-purpose hardware to provide for the distribution of visual material. An extensive review of such systems may be found in Watanabe et al. [85]. Increasingly, several media have been combined into a single system, e.g., audio with handwriting (Bordewijk [78]), audio with computer generated text and data (Chatterley et al. [86]), sometimes extended with freeze frame video (Tsuruta et al. [86]). Some systems combine motion video with traditional media (INVITE-64, Judice & LeGall [87]). Others include audio, data, and motion video into one terminal (Sciarappa [89], Clark [92]). Various systems are reported using ISDN connections (Chatterley et al. [86], Tsuruta et al. [86], Sciarappa [89], Clark [92]), while others are based on LAN communications (Sakata & Ueda [90], Ahuja & Ensor [92]), or are more network independent (Robinson et al. [91]). The majority of the most recent publications describe developments based on workstations or PCs and accompanying peripherals. This approach leads to an integrated environment for real-time conferencing, networking, personal computing and CSCW applications. However, most systems require special-purpose hardware, except Sakata & Ueda [90] and Robinson et al. [91]. Both these latter systems are based on PCs, and provide real-time multipoint conferencing services. Communication between multiple users over common media such as an audio channel and a shared screen, requires well-developed floor control management. Floor control protocols have been discussed by several authors (Aguilar et al. [86], Sakata & Ueda [90], Robinson et al. [91], Clark [92]).
Some authors describing the development of a real-time multimedia conferencing terminal are considering the possibility of tele-education applications (e.g., Ahuja & Ensor [92], Robinson et al. [91]). However such systems are primarily developed for conferencing and CSCW purposes. Further it can be concluded that most systems are not suitable for narrowband communication and/or require special purpose hardware.

By contrast, the terminal described in this chapter is PC-based and does not require dedicated hardware. Besides, the system has built-in flexibility with respect to the available bandwidth and operates well at 9600 bit/s. Moreover, it was designed to enable efficient tele-teaching of multiple remote groups, meeting the requirements stated above.

Fig. 8.1 Tele-education terminal hardware configuration

8.4 System overview

The system designed enables a teacher to provide real-time courses to a number of small remote groups. This may be in combination with a live class, making efficient use of the available teaching resources. The diagram of Fig. 8.1 shows a point-to-point terminal configuration. The system is based on an IBM PC/AT-compatible personal computer (with either an 80286, 80386 or an 80486 processor), with display, communications and input hardware. Input media available are text and graphics via keyboard and mouse; handwritten information using an electronic writing tablet; raster-scanned documents.
8.4 System overview

Inverting Amplifier

\[ \begin{align*}
I_i &= \frac{V_i}{R1} \\
V_o &= -I_i \cdot R2
\end{align*} \quad \text{Av} = \frac{V_0}{V_i} \quad R_i = \frac{V_i}{I_i} = R1
\]

*Fig. 8.2 Example of shared-screen area, LNV (small size), and the control menu at the top of the screen with a dialog box popped up

(document camera or scanner); and a video-camera to relay, e.g., the image of the teacher. One display is utilized, with a resolution of 640 x 480 pixels and at least 16 colours (VGA standard). Two-way voice communication is provided by standard audio equipment. Two PSTN circuits are employed to exchange audio and data. Another option is the application of a dedicated satellite network to broadcast a multiplexed audio and data signal, with the possibility of phone-in feedback via the PSTN.

The teacher may use the system off-line (i.e., without active communication to remote groups) for the production, management and archiving of educational materials. All off-line functions can be executed in on-line mode, too, thus providing maximal flexibility regarding the presentation of visuals during a lecture. Moreover, the system can be used in a multipoint configuration and supports effective floor control.

The software should be modular. This ensures that the addition of new media or new communication interfaces, changes in the user interface or in the screen resolution, or other modifications, would have limited effect on the software. As shown in Fig. 8.3, six main groups of software modules are distinguished: control, display interface, com-
Fig. 8.3 Software architecture showing the six groups of modules and the major data flows

Communications, file & macro management, multimedia processing and user interface. The main function of the control group is to provide a non-preemptive task scheduler for the various communicating processes. Other functions are peripheral configuration control, debug information (only for development purposes), and error functions (peripheral operation, memory and storage device capacity, communication links, etc.). The second group of modules provide the display management and driver functions including window control, cursor and pointer generation, and dialog box features. Display standards supported include the EGA standard, the VGA standard (16 colours; high resolution), and extended VGA (256 colours; high resolution). Communications software resides in a third group of modules consisting of application-level send and receive processes and a
multipoint communication control layer. Additionally, this group contains modules for the management and execution of pretransfer functions. All user actions are interpreted by the user-interface module, which makes appropriate calls to the other modules if applicable. An extensive help utility, multipoint floor control, and various session monitor and directory services belong to this group of modules. The group of multimedia processing modules include the handling of LFV and the input and editing of text, structured graphics, raster-scanned graphics (document camera), and handwriting. Besides, the current virtual shared-screen buffer (SSB) is controlled by these modules. The final group of modules encompasses presentation file management and macro file management functions.

8.5 Multimedia processing

8.5.1 Page structure

The system supports the following data-types for display at the shared-screen area: text, structured graphics, linedrawing (i.e., handwritten graphics), and raster-scanned pictures. Processing and transmission functions of more complex data-types like video, audio or animation tracks are not integrated due to three reasons: 1) these media are relatively less used in day-to-day lecturing; 2) the large data sizes would result in unacceptable transfer times, even if adequate compression methods would be applied; 3) these media types would require specific and complicated editing software. A viable approach may be to produce and edit such media off-line, by off-the-shelf multimedia tools and authorware. Then transfer time problems may be overcome by using a database synchronization technique (Bruyne [91]).

The organization of data generation, manipulation and storage should be object-oriented. This allows every object to be separately modifiable and to be moved to any arbitrary location at the screen. Moreover, by the application of hierarchical structures, higher-order objects can be produced and edited, composed of several basic objects. Further, an object-oriented design allows for media-specific compression methods and for the application of semantic and media structure techniques, in order to reduce end-to-end transfer times (Rosenberg et al. [92]).

The shared-screen window may display pages consisting of various objects on a one-colour background. To allow fast editing and re-display operations, an object-oriented representation of the shared-screen contents is kept in memory (shared-screen buffer,
8.5.2 Text and Graphics

The text editor enables text entry at typing speed. A cursor indicates the position of the next character, and may be moved to any place on screen by mouse or arrow-key actions. A limited number of fonts (12) and colours (16) can be selected. For tele-education purposes a more extensive set of these attributes is less functional than other applications as, e.g., desk-top publishing. On the contrary, this would contribute to an unnecessary complicated user interface (UI). Consequently, the amount of attribute selections has been limited for all data-types. In this context, three character sizes may be used, corresponding to 20, 10 and 5 lines per screen, respectively. Smaller sizes suitable for desk-top applications are not appropriate for class situations, taking into account practical viewing distances. As each character is a single object, it may be represented in its own colour, font, and size. During text entry basic editing functions like backspace, delete, insert, etc., are possible, and the current font, colour and size are displayed at the menu bar. Besides move, copy, erase and modify functions, also a clear-screen (only affecting text) and a back operation is available. Further, editing of text objects may be on a single object or on a rectangular field basis.

The graphics editor supports generation and editing of a limited set of structured objects (i.e., line, bar, rectangle, circle, disc, segment, arc, ellipse, filled ellipse, polygon, filled polygon). Besides, frequently used higher-order objects like a line-graph and a pie-chart are available. The user is prompted to input the number of slices or points via a dialog box, followed by the corresponding values. The available editing functions include copy, erase, move, modify, zoom, rotate, mirror, and clear-screen. Selectable attributes are colour, line-thickness, line-style, and fill-pattern. Further a line-grid or a point-grid may be put on screen to facilitate accurate drawing.

8.5.3 Raster-scanned images

Using a scanner-device, visuals of paper-sheets may be digitized and transmitted. A video-camera mounted on a vertical stand, together with a frame-grabber PC-board, enables fast acquisition of these type of images, with sufficient resolution. The system operates in either of the two modes: 256-colours pictures or bi-level pictures. The former may be used to transfer any image of a paper-sheet, while the latter enables the system to be used as a 'tele-overhead projector'. The teacher simply puts a sheet on the document-stand, followed by a keystroke triggering the capturing, digitization, compression and transmission process. In other words, in this way the system can be used by a novice user.
(e.g., a guest lecturer). To simplify system design, the image acquisition and compression process in the 256-colour mode is equal to the LFV process (see 8.5.6), characterized by fast and effective software-based compression algorithms. The highest resolution is 512 x 384 pixels, corresponding to 80% of the VGA-screen. In the bi-level mode (any combination of two colours may be selected) a basic run-length compression scheme is implemented (Jayant & Noll [84]). Other algorithms such as CCITT Group IV facsimile coding might result in higher compression ratios, but would slow down system response-times due to a higher computational load. The usual set of editing functions (refer to 8.5.2) is extended in this case with clip, declip, and zoom possibilities. Further, image parameters like contrast, brightness, and threshold can be adjusted.

8.5.4 Linedrawing

Linedrawing may be input using an electronic writing tablet with a handheld stylus (Kegel & Bons [77]). A number of attributes (line-style, line-thickness, colour) can be selected.
The same set of editing functions is supported as with structured graphics. The data compression method used is Modified Tunstall Coding (MTC - see Chapter 5) which is an optimized form of Differential Chain Coding (DCC). This involves the application of an extended set of codewords representing segments of relative or absolute vectors. A significant reduction of the coding rate results, as compared to DCC. This efficient method preserves DCC's typical robustness against transmission errors. Further the system is equipped with a fill facility in such a way that graphical editing functions remain feasible. The additional coding information consists of only 3 elements, to be added to the DCC trace header.

The extension or editing of on-screen-displayed objects by means of an electronic writing tablet implies an eye-to-hand coordination problem. Therefore a Relative Coordinate System (RCS) facility has been implemented by which the user is able to position handwritten graphics on screen at a precisely defined location. For a detailed discussion on compression, filling, and RCS issues refer to Chapter 6.

8.5.5 Macro libraries

The user is able to define any object or combination of objects as a macro. Any group of macros, called a macro library, can be selected for retrieval of existing macros or to add new defined macros. The file organization of a macro library is equal to that of a presentation. Further it is possible to display the directory of a selected library containing macro names together with short descriptive strings. The latter may be input in the macro definition phase. The macro facility is of importance to the efficient production of pages. For example, lecturing in electronics involves frequent use of electronic circuits drawings composed of standard component symbols. In this case, an electronics symbols library will effectively reduce the production time of a page. Each discipline requires its own typical macro libraries. Libraries can be easily exchanged and distributed due to their DOS-file compatible organization.

8.5.6 Low Frame-rate Video (LFV)

The requirements include a full-duplex visual communication component. The students should be able to view the teacher by, e.g., a head-shoulder image. At the same time the teacher should be able to watch the class situation. Often, small-sized images may be sufficient, inserted at a corner of the screen. However at any time the teacher should be able to switch to an image of high resolution (i.e., full screen) for a closer inspection. On the other hand, some situations require a full-screen image to be presented to the students,
multipoint communication control layer. Additionally, this group contains modules for the management and execution of pretransfer functions. All user actions are interpreted by the user-interface module, which makes appropriate calls to the other modules if applicable. An extensive help utility, multipoint floor control, and various session monitor and directory services belong to this group of modules. The group of multimedia processing modules include the handling of LFV and the input and editing of text, structured graphics, raster-scanned graphics (document camera), and handwriting. Besides, the current virtual shared-screen buffer (SSB) is controlled by these modules. The final group of modules encompasses presentation file management and macro file management functions.

8.5 Multimedia processing

8.5.1 Page structure

The system supports the following data-types for display at the shared-screen area: text, structured graphics, linedrawing (i.e., handwritten graphics), and raster-scanned pictures. Processing and transmission functions of more complex data-types like video, audio or animation tracks are not integrated due to three reasons: 1) these media are relatively less used in day-to-day lecturing; 2) the large data sizes would result in unacceptable transfer times, even if adequate compression methods would be applied; 3) these media types would require specific and complicated editing software. A viable approach may be to produce and edit such media off-line, by off-the-shelf multimedia tools and authorware. Then transfer time problems may be overcome by using a database synchronization technique (Bruyne [91]).

The organization of data generation, manipulation and storage should be object-oriented. This allows every object to be separately modifiable and to be moved to any arbitrary location at the screen. Moreover, by the application of hierarchical structures, higher-order objects can be produced and edited, composed of several basic objects. Further, an object-oriented design allows for media-specific compression methods and for the application of semantic and media structure techniques, in order to reduce end-to-end transfer times (Rosenberg et al. [92]).

The shared-screen window may display pages consisting of various objects on a one-colour background. To allow fast editing and re-display operations, an object-oriented representation of the shared-screen contents is kept in memory (shared-screen buffer,
SSB). Fig. 8.4 shows the different data formats corresponding to the four data types allowed. The structure of the SSB is depicted in Fig. 8.5, showing two of the three components: 1) a header containing the page identifier, the specification of the background-colour, and the comment label; 2) an object pointer list (objects-list); 3) the actual object data representations (objects-buffer). The comment label is an optional character-string which can be attached to a particular page for recognition purposes. The objects-list guarantees fast access, even to a large-sized objects-buffer. The contents of the objects-buffer are medium-compressed in such a way that object-oriented editing remains
possible, but that no additional compression is necessary for file storage and transmission purposes. The order of the list corresponds to the progression of object drawing on screen. The author should be able to change this order for various reasons. First, 'early' input objects overlapped by 'later' objects might be transferred to the foreground. Second, in compound pages the display-order of objects of different data-types should correspond to the way a user interprets visual information. This is especially important if displaying objects of a particular data type takes a perceivable amount of time as with linedrawings. For example, if a geographical map containing text-based (i.e., typographic) location names with linedrawing-based map contours is to be displayed, then a display progression of text followed by the map contours seems not logical to the user and may be confusing. On the other hand, e.g., handwritten annotations to a text-based page should be displayed in the opposite order. Therefore a back function is added to the different editors to allow an object or a group of objects to be transferred to the background.

Objects located in a user-selectable rectangular field on screen may be edited independent of the particular data types, by a general editor (functions: move, copy, erase, rotate, mirror). Further, four editors are included operating on objects of a specific data-type: a text editor, a structured graphics editor, a raster-scanned picture editor, and a linedrawing editor. This allows specific data-type dependent editing functions and easy manipulation of structures composed of multiple objects of a single data-type without affecting other data-type objects.

The fifth data-type used is Low Frame-rate Video (LFV) with its own resizable window. The purpose of LFV is to provide a form of real-time video communication. Consequently, LFV data cannot be edited, stored or retrieved like the shared-screen objects.

A page can be stored in the current DOS directory as a separate file. The filename prefix defines the current presentation, and the suffix specifies the sequential page number. Hence any number of pages (but ≤ 100) may be grouped together as a presentation with a user-specified name. Typically, such a presentation will be presented in sequential order. However the system allows the teacher to choose any arbitrary succession. Consequently, the next or the previous page in a presentation can be retrieved by a single keystroke. Any other page should be specified by its number. The system supports deletion and insertion of pages and it is possible to display the directory of a presentation. The objects of any page can be added to the current page by an overlay function. Other file management functions are: change of presentation name, and display or change DOS directory.
8.5.2 Text and Graphics

The text editor enables text entry at typing speed. A cursor indicates the position of the next character, and may be moved to any place on screen by mouse or arrow-key actions. A limited number of fonts (12) and colours (16) can be selected. For tele-education purposes a more extensive set of these attributes is less functional than other applications as, e.g., desk-top publishing. On the contrary, this would contribute to an unnecessary complicated user interface (UI). Consequently, the amount of attribute selections has been limited for all data-types. In this context, three character sizes may be used, corresponding to 20, 10 and 5 lines per screen, respectively. Smaller sizes suitable for desk-top applications are not appropriate for class situations, taking into account practical viewing distances. As each character is a single object, it may be represented in its own colour, font, and size. During text entry basic editing functions like backspace, delete, insert, etc., are possible, and the current font, colour and size are displayed at the menu bar. Besides move, copy, erase and modify functions, also a clear-screen (only affecting text) and a back operation is available. Further, editing of text objects may be on a single object or on a rectangular field basis.

The graphics editor supports generation and editing of a limited set of structured objects (i.e., line, bar, rectangle, circle, disc, segment, arc, ellipse, filled ellipse, polygon, filled polygon). Besides, frequently used higher-order objects like a line-graph and a pie-chart are available. The user is prompted to input the number of slices or points via a dialog box, followed by the corresponding values. The available editing functions include copy, erase, move, modify, zoom, rotate, mirror, and clear-screen. Selectable attributes are colour, line-thickness, line-style, and fill-pattern. Further a line-grid or a point-grid may be put on screen to facilitate accurate drawing.

8.5.3 Raster-scanned images

Using a scanner-device, visuals of paper-sheets may be digitized and transmitted. A video-camera mounted on a vertical stand, together with a frame-grabber PC-board, enables fast acquisition of these type of images, with sufficient resolution. The system operates in either of the two modes: 256-colours pictures or bi-level pictures. The former may be used to transfer any image of a paper-sheet, while the latter enables the system to be used as a 'tele-overhead projector'. The teacher simply puts a sheet on the document-stand, followed by a keystroke triggering the capturing, digitization, compression and transmission process. In other words, in this way the system can be used by a novice user
(e.g., a guest lecturer). To simplify system design, the image acquisition and compression process in the 256-colour mode is equal to the LFV process (see 8.5.6), characterized by fast and effective software-based compression algorithms. The highest resolution is 512 x 384 pixels, corresponding to 80% of the VGA-screen. In the bi-level mode (any combination of two colours may be selected) a basic run-length compression scheme is implemented (Jayant & Noll [84]). Other algorithms such as CCITT Group IV facsimile coding might result in higher compression ratios, but would slow down system response-times due to a higher computational load. The usual set of editing functions (refer to 8.5.2) is extended in this case with clip, declip, and zoom possibilities. Further, image parameters like contrast, brightness, and threshold can be adjusted.

8.5.4 Linedrawing

Linedrawing may be input using an electronic writing tablet with a handheld stylus (Kegel & Bons [77]). A number of attributes (line-style, line-thickness, colour) can be selected.
The same set of editing functions is supported as with structured graphics. The data compression method used is Modified Tunstall Coding (MTC - see Chapter 5) which is an optimized form of Differential Chain Coding (DCC). This involves the application of an extended set of codewords representing segments of relative or absolute vectors. A significant reduction of the coding rate results, as compared to DCC. This efficient method preserves DCC's typical robustness against transmission errors. Further the system is equipped with a fill facility in such a way that graphical editing functions remain feasible. The additional coding information consists of only 3 elements, to be added to the DCC trace header.

The extension or editing of on-screen-displayed objects by means of an electronic writing tablet implies an eye-to-hand coordination problem. Therefore a Relative Coordinate System (RCS) facility has been implemented by which the user is able to position handwritten graphics on screen at a precisely defined location. For a detailed discussion on compression, filling, and RCS issues refer to Chapter 6.

8.5.5 Macro libraries

The user is able to define any object or combination of objects as a macro. Any group of macros, called a macro library, can be selected for retrieval of existing macros or to add new defined macros. The file organization of a macro library is equal to that of a presentation. Further it is possible to display the directory of a selected library containing macro names together with short descriptive strings. The latter may be input in the macro definition phase. The macro facility is of importance to the efficient production of pages. For example, lecturing in electronics involves frequent use of electronic circuits drawings composed of standard component symbols. In this case, an electronics symbols library will effectively reduce the production time of a page. Each discipline requires its own typical macro libraries. Libraries can be easily exchanged and distributed due to their DOS-file compatible organization.

8.5.6 Low Frame-rate Video (LFV)

The requirements include a full-duplex visual communication component. The students should be able to view the teacher by, e.g., a head-shoulder image. At the same time the teacher should be able to watch the class situation. Often, small-sized images may be sufficient, inserted at a corner of the screen. However at any time the teacher should be able to switch to an image of high resolution (i.e., full screen) for a closer inspection. On the other hand, some situations require a full-screen image to be presented to the students,
for instance to transfer the image of a blackboard. Therefore the system supports three image sizes: 128 x 96 pixels, 256 x 192 pixels, and 512 x 384 pixels, corresponding to 20, 40, and 80 percent of the full-screen dimensions, respectively (see Fig. 3.3). These video images are captured and digitized periodically by a frame-grabber residing in the PC (add-on board). As described in Chapter 4, a software-based image compression method has been developed, combining a high compression ratio with a low CPU-load, called FEVQ (Fast and Efficient Vector Quantization). This algorithm applies LUT-based MSVQ on 2x2 pixel blocks in combination with DPCM on the block means and Modified Tunstall Coding. Further, low complexity interframe compression has been realized by conditional replenishment. A two level approach was adopted: replenishment of 2x2 pixel blocks and of 128x96 pixel blocks (called: sub-images). With sub-image replenishment based on a sample check mechanism, a very efficient encoder implementation has been implemented. The encoding process is appropriate for extension to colour encoding. Furthermore it can easily be re-ordered into a two stage progressive transmission scheme to reduce the effective transfer time by 50%. Besides activation, de-activation, and image-size switching, the user is able to control image contrast and brightness.

8.6 Communication services

The system facilitates multipoint sessions in a primary/secondary mode and point-to-point sessions in a peer-to-peer mode. The multipoint communications aimed at real-time tele-education includes multiple call set-up, and adequate floor control mechanisms. Further we discuss an application protocol to reduce transfer times, a number of additional services, and some data-link layer issues.

![Diagram of terrestrial multipoint configuration using the local PSTN](image)

*Fig. 8.7 Terrestrial multipoint configuration using the local PSTN*
8.6.1 Terrestrial and satellite-based configurations

As has been discussed in 3.3, the main information stream from teacher to class should be multicast to all remote groups. This may be via a terrestrial public network (PSTN) using separate audio and data circuits, or by satellite communication where a multiplexed voice-data signal would be the most appropriate choice. For both types of situations, data-rates of at most about 9600 bit/s apply. The PSTN solution may involve continuous feedback from all remote sites including voice and data. However in areas where the satellite option is preferred, phone-in feedback will probably be the optimal solution. This may involve multiplexed voice and (low bitrate) data using a dial-up PSTN link. The description in the remaining part of this section will be devoted to such a terrestrial configuration.

The audio signals require a bridge function, which may be delivered as a PTT service. However in situations where multiple groups are taught simultaneously, a requirement exists for audio-volume control. The teacher should be able to suppress or present any of the audio signals from the different groups. Hence the bridge function should be implemented using in-house equipment (refer to Fig. 8.7). A problem inherent to audio transmission of conferences is the acoustic coupling between the loudspeaker and the microphone. These couplings at both terminal rooms create an acoustic loop. Once the

![Diagram of satellite-based broadcast configuration using the PSTN for phone-in feedback]

*Fig. 8.8 Satellite-based broadcast configuration using the PSTN for phone-in feedback*
loop gain for the voice signals becomes greater than 1, howling tends to occur. One method to prevent this is the application of a voice switch which enforces transmission of the voice signal in one direction at a time (Watanabe et al. [85]). All groups provide separate data feedback. At the primary station each of these signals is continuously received, deformatted and being made available to the application level entities. However the actual use of the data (e.g., displaying of LFV data in a window on the screen) depends on the current floor control mode.

8.6.2 Floor control and call set-up

An effective floor-control mechanism for tele-education purposes differs from concepts developed for general tele-conferencing applications. One of the primary objectives is to approximate face-to-face teaching as close as possible. Hence the interaction process involving simultaneous teaching of multiple remote groups by one teacher should be equal.
to those apparent in a live class. Floor control encompasses admission of shared-screen input, and video and audio distribution over the various groups. The complete process should be managed and controlled by the teacher from the primary station. The discourse mode involves the remote classes receiving the teacher's audio and video signals (LFV), and having no active access to the shared screen. However, at any time, the teacher may select any of the connected groups for interaction (interaction mode). In this mode the group's audio-signal is bridged together with the teacher's audio signal and relayed to all sites. Further, at all other sites including the primary station, the group's video signal will be displayed. Third, at the station of the group, input to the shared screen is allowed via keyboard, mouse, writing tablet or document scanner. The teacher is always able to inhibit or erase these remote contributions. Any of the groups may request attention of the teacher by an interrupt request function. Activation triggers an audio-visual alarm at the primary's UI, including the display of a group identification. If the teacher grants the request (by a single keystroke action) then automatically the interaction mode is entered. A resettable attention-request mechanism is also available to the teacher in order to notify any or all remote classes. Table 8.1 lists the various floor control functions as available at the primary station. Fig. 8.9 shows how the transitions between these various modes can be effectuated through simple UI actions.

Note that this solution is easily extended with a peer-to-peer mode, as applicable in a point-to-point configuration. If one remote group is connected to the teacher's station, then the 'interaction mode' can be used for this purpose, provided that the 'primary' station is inhibited to overrule the shared-screen actions of the 'secondary' in this case. However the system is not designed for multipoint conferencing in a peer-to-peer mode. This would require a more complicated video data distribution scheme in order to provide every conferee with the images of all others (compare Robinson [91]). Though, in practical situations, a form of multipoint conferencing may be possible using the primary/secondary concept as explained above. In that case the user of the primary station should act as the session chairman.

The system has special built-in call set-up features. The normal procedure is that the remote user calls the primary station at the time scheduled for a particular lecture he wants to follow. This combined data and audio call set-up is software controlled including automatic redialing, user notification, and telephone number programmability. Thus the primary user should only bring his system in an accept mode, and wait for the establishment of the various connections. Obviously, the primary station reports connection-status transitions to the user.
Table 8.1  Floor control modes for multipoint tele-education with two remote groups

<table>
<thead>
<tr>
<th>Mode</th>
<th>Origin of audio at all sites</th>
<th>Origin of video image on monitor of</th>
<th>Access to shared screen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Teacher (T)</td>
<td>Group A</td>
</tr>
<tr>
<td>Discourse Interaction</td>
<td>T</td>
<td>A or B</td>
<td>T</td>
</tr>
<tr>
<td>Interaction</td>
<td>A+T</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>Interaction</td>
<td>B+T</td>
<td>B</td>
<td>B</td>
</tr>
</tbody>
</table>

8.6.3 Pretransfer protocols

The following techniques have been applied to minimize page transfer-times. First, the allowed set of data-types does not involve objects of extreme data size. Second, media-compression is used for raster-scanned objects and linedrawing data. Third, semantic and media structuring have been applied (Rosenberg et al. [92]). Semantic structure uses author specifications to eliminate redundant transmission. A page may consist of a fixed 'background' containing various objects upon which other objects appear following proceed commands or even move or disappear by on-line editing commands. The background is sent only once. The display machine understands the semantics of the background and will repaint parts of it when necessary, without the need for communication with the sender. The media structure reduces the amount of information transmitted by using the structure of specific media. This applies to structured graphics and text. Instead of transmitting raster-scanned objects, structural information is sent: graphics type identification, text characters, fonts, coordinates, other attributes, etc.

All three methods aim at compressing the page data, resulting in shorter transmission times and efficient data storage. However, real-time tele-education requires any page to

Table 8.2  Typical page transfer times with and without pretransfer mechanism

<table>
<thead>
<tr>
<th>Page</th>
<th>kbyte</th>
<th>No pretransfer</th>
<th>Pretransfer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1.2 kbit/s</td>
<td>9.6 kbit/s</td>
</tr>
<tr>
<td>Text &amp; Graphics</td>
<td>10</td>
<td>74</td>
<td>9</td>
</tr>
<tr>
<td>Handwriting</td>
<td>10</td>
<td>74</td>
<td>9</td>
</tr>
<tr>
<td>Bi-level picture</td>
<td>8</td>
<td>59</td>
<td>7.3</td>
</tr>
<tr>
<td>16-level picture</td>
<td>25</td>
<td>184</td>
<td>23</td>
</tr>
<tr>
<td>Compound page</td>
<td>50</td>
<td>367</td>
<td>46</td>
</tr>
</tbody>
</table>
be transferred in a few seconds; this cannot be achieved for all types of pages in the case of narrow bandwidths. Therefore a time shifting method has been applied, referred to as pretransfer, which exploits periods of low channel utilisation during the session for automatic transmission of page sequences in background. When the lecture is in progress, arbitrary objects and pages can be added on-line. The pretransfer process is not noticeable by the user and results in virtually zero transfer times (VZT). The pretransfer process starts immediately when the connection is established. Secondary stations are protected from disk overload. Insufficient storage capacity results in a message to the user of the primary station and then the pretransfer mechanism stops. After an accidental link dropout, the process resumes at the last point before the dropout. If the primary retrieves a page for display, the pretransfer process checks its status information and decides if the complete page or only a copy of the retrieval command should be sent. In the first case, the primary user is notified on completion of the page transmission process. An extensive description of point-to-point and multipoint protocols can be found in Chapter 7.

The figures of Table 8.2 illustrate the efficiency of the pretransfer method. The end-to-end transfer times of various pages are compared to situations without a pretransfer mechanism. The data sizes are typical values for rather 'crowded' pages, and a 10 percent framing overhead is taken into account.

8.6.4 Additional services

By default, the system enters the interaction mode at start-up, meaning that a peer’s input actions to the shared screen cannot be overruled. In a point-to-point configuration, both stations are in this interaction mode and floor control is subject to social rules only. The primary mode can only be entered via a password protected log-in procedure. In discourse mode a station is able to control the secondary stations to a substantial extent as has been discussed in 8.6.2. Moreover, only a primary station can set-up multipoint connections.

During the session, part of the menu-bar at the top of the screen is allocated to the continuous display of some communication parameters. These include icons indicating the calling progress (no connection / calling / connection established) with each of the connected stations. Another indicator provides information whether data is being received, transmitted or neither of these modes. Third, a transfer-time display has been incorporated. As soon as the user sends a new page, a down-counter starts running indicating the number of seconds left to the instant the page will be displayed at the distant screen(s). This time is calculated from the data size of the page and the bit rate of the slowest connection. Note that this feature is not relevant when page transfer times are
virtually zero due to the pretransfer protocol. Besides these continuously displayed parameters, the user may pop a monitor window up, showing relevant information on the various connected links like effective bit rate, bit error rate (BER), and elapsed connection time.

By the align command the primary user is able to re-synchronize the shared-screen areas and relevant system parameters of the various remote stations. This command is issued automatically at connection establishment. In specific circumstances (poor lines, or terminal equipment problems) the manual version of the command may be useful.

8.6.5 Data-link layer issues

The lower layers of the OSI reference model are not a subject of this study. Still, a few issues should be discussed: error control, flow control, and priority management. Obviously, the network may have heterogeneous properties regarding bandwidths, bit error rates and CPU performances. Hence, effective flow control mechanisms are necessary both at application level and data-link level. Further, the data link should provide a full-duplex data channel with an adequate ARQ or FEC mechanism to reduce the effects of transmission bit errors down to an acceptable level. A data priority scheme is also required. Six types of data are distinguished according to descending priority:

1) System commands (change page; scroll; zoom screen; scratch window; etc.)
2) Pointer and marker data
3) On-line entered objects
4) Pages
5) LFV data
6) Pretransfer data

It is of vital importance that the first two types of data take absolute priority over the others. Further it is observed that on-line entered objects and pages are directly related to the current progress of the lecture and should have precedence above LFV and pretransfer data. Pretransfer should not be completely suppressed if the remaining bit rate capacity would be insufficient for normal execution of LFV. In that case a part of the capacity is still to be allocated to pretransfer. Hence the transport layer should provide the possibility to establish a prioritized set of connections.
Table 8.3  Memory occupation (excluding LFV buffers) in kBytes.

<table>
<thead>
<tr>
<th>Application code</th>
<th>150 kB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application data area</td>
<td>50 kB</td>
</tr>
<tr>
<td>Drivers</td>
<td>50 kB</td>
</tr>
<tr>
<td>Codec tables</td>
<td>80 kB</td>
</tr>
<tr>
<td>Shared-Screen Buffer</td>
<td>50 kB</td>
</tr>
<tr>
<td>Window Manag. buffers</td>
<td>60-120 kB</td>
</tr>
<tr>
<td>Total</td>
<td>440-500 kB</td>
</tr>
</tbody>
</table>

8.6.6 LFV memory occupation

The PC-memory budget is a critical issue in the implementation of the described system. Above all, the LFV codec buffers tend to occupy a substantial amount of memory. Table 8.3 lists the memory occupation for the application code, and the various buffers and tables, but excluding the LFV buffers. From this list it is concluded that the amount of memory allocated to the LFV buffers should be $\ll 200$ kByte.

Obviously, in a multipoint configuration the memory budget of the primary station will be the most critical one. Therefore let us consider the amount of LFV buffer memory in a primary station, denoted by $M$. Let $s$ and $p$ denote the LFV image size as transmitted by the secondary and the primary station, respectively, with $s, p \in \{0,1,2\}$. Let $L_s$ denote the number of pixels of a small-size LFV image (i.e., $128 \times 96 = 12288$ pixels). In the primary station two LFV buffer types exist. One buffer type is used for conditional replenishment coding. Each group of 4 pixels ('subblock') occupies eight bits, representing the subblock mean and colour values. The other buffer type is used in the decoding process corresponding to 16 bits per subblock. The reader is referred to Chapter 4 for a detailed description. Let there be $N$ secondary stations. Hence $N$ buffers should be kept in memory to allow fast switching between the various LFV images in order to adequately support inspection and interaction purposes. It is easily verified that $M$ becomes

$$M = L_s(2^{2p-2} + 2^{2s-1}N)$$

(8.1)

If $p = s = 2$, $N$ must be less than 2 to meet the above requirement. Hence in multipoint configurations $s$ should be limited to 1 (medium size) if $p = 2$ (large).
Three alternative solutions may be distinguished to handle multipoint LFV with \( p = 2 \) and \( s = 1 \).

1) All LFV images are continuously decoded and stored. Hence switching over from one secondary station to the other does not take time. The memory occupation is given by eqn. (8.1).

2) Only the LFV data of one secondary station at a time is stored and processed. This minimizes the memory requirements. \( M \) is equal to

\[
M = L_s (2^{2p-2} + 2^{2s-1})
\]  

(8.2)

However, switching over to the LFV image of another station requires the transmission of a complete image, without conditional replenishment. The corresponding transmission time \( T \) is

\[
T = L_s \frac{2^{2s}R}{B}
\]  

(8.3)

Here \( R \) is the average number of bit/pel of the compressed image representation.

3) Only the LFV image currently displayed is continuously decoded. The others are received and stored, without decoding. This saves memory space, but switching takes a short amount of time due to the decoding process. \( M \) becomes:

\[
M = L_s (2^{2p-2} + 2^{2s-1} + \frac{R(N-1)}{4}2^{2s-1})
\]  

(8.4)

Fig. 8.10 Memory occupation of LFV buffers in primary station as a function of \( N \) (image size of secondary station \( s=1 \), and of primary station \( p=2 \), channel data rate \( B=9600 \) bit/s, \( R=1 \) bit/pel)
Fig. 8.10 presents $M$ for the three solutions as a function of $N$, for $s=1$, $p=2$, $R=1$, $B = 9600$ b/s, and $L_s = 12288$. Note that for solution 2) the transmission time $T$ equals 5.1 s. It is concluded that solution 3) prevails by the combination of efficient memory use with a short switch-over time, even at large $N$ values.

### 8.7 The user’s view of the system

#### 8.7.1 General

Fig. 8.2 and 8.6 show the user’s view of the system on a VGA display: a large shared space (95% of the screen area), with control menus along the top. Mouse actions or keystrokes affectuate menu items which can include activation of a deeper or a higher menu. The number of menu levels is limited to three (for a few items to four). Dialog boxes are used for detailed user input as necessary for certain commands. System information such as help, directories, and status data is communicated to the user through windows. Although windows and dialog boxes are overlayed to the shared space, they never obstruct the execution of editing commands. The experienced user has the possibility to switch rapidly between the items of the various sub-menus using shortcuts (function keys) instead of the menu interface for the most frequently used commands. In the menu bar, icons and status display items inform on current attributes, menu status, communication modes, presentation name, etc. Appendix B gives the complete menu structure, including all the commands available. The shared space can be used as a (enhanced) whiteboard, overhead projector, slide projector, or a combination of these. This area may be overlayed by LFV images of various size as shown in Fig. 3.3.

#### 8.7.2 Shared-screen manipulation

As discussed in 3.2.2.4, a scrolling and zooming mechanism has been implemented. The default visible part of the shared screen (the view port) corresponds to the upper left quadrant of the entire page area. Horizontal and vertical scrolling is possible by steps of 1/4 of the total width and height, respectively. Further, the viewport can be zoomed out in two steps to maximal coverage of the entire area. This feature is useful if, e.g., an extensive diagram or a lengthy derivation has to be presented. The overview and the logical cohesion of the various components can be comprehended more easily by adequate
use of these scrolling and zooming mechanisms. A micro-image of the shared area at the menu bar with the current location and size of the viewport supports these functions.

Another useful facility is the *scratch* window. At any time during the presentation this window may be popped up and used for inputting and editing of information similar to the normal shared-screen window, but with the exception of storage to a presentation file. However this scratch information remains retrievable until the session is closed. The feature allows the teacher to give improvised 'intermezzo' explanations, without affecting the preprogrammed materials.

A *pointer* (with two possible orientations) can be popped up and moved around the shared screen. The changing position of the primary's pointer is continuously transmitted to the secondary stations in order to maintain a synchronous copy at all screens. In the interaction mode the secondary user is also able to control a pointer which is visible at all sites. However, the latter is automatically disabled if the primary's pointer has been activated. A *marker* function allows the teacher to highlight any part of the screen by a selectable colour by using the mouse as a painting device. Two 'line-widths' are possible. Like in the case of pointing, marker information is continuously sent to all connected stations, and may also be input by a secondary user during the interaction mode. Marker information cannot be stored and so disappears at the next page transition.

During page preparation *pause marks* may be inserted between arbitrary objects of a page. When such a page is called up for presentation later on, the displaying process of the consecutive objects stops at each pause mark and is resumed again by a *proceed* command. This feature allows the progressive introduction of preconceived material. Pause-marks are stored and manipulated like other object types (see Fig. 8.5).

### 8.7.3 Course of the session

For a rather elementary level, only a few training instructions enable the user (both teacher and students) to operate the system effectively. As soon as the multipoint connection has been set-up, the teacher may suffice to put visuals on the stand of the document camera. This is comparable with using an overhead projector in conventional teaching. At 9600 bit/s the transfer-times of these visuals in general are not more than 5 seconds due to run-length coding of two-level digitized images. Using a small-size LFV image inserted at a corner of the screen the students view the teacher, and vice versa. Further, two-way audio communication is available. In other words, this low-level way of operating the system just requires putting sheets on the document stand followed by keystrokes to trigger the sheet transmission process.
For on-line inputting of text or drawings the electronic writing tablet is available. A pointer can be moved around the shared screen using a mouse device. When a student wants to raise a question or make a statement, this may be interrogated by an interrupt request. The result is an audio-visual indication at the teacher's UI. By only one UI action the floor is passed to the student making the request. Then the LFV image of the student appears on all the screens and his audio signal, bridged with the teacher's, is distributed to all the sites as well. Moreover, the student gets access to the shared screen and may input text, graphics and handwriting. In this way the class interaction process is closely approximated, requiring only a minimal number of UI actions.

At the more experienced user level, numerous useful additional features are available, as has been described above. The teacher is able to prepare structured presentations composed of compound pages prior to the session. For this purpose an extensive set of editing and filing functions may be used. These pages may be edited and/or supplemented on-line, during the session. The effective page transfer times are negligible due to the pretransfer mechanism, even in the case of high resolution pictures. By LFV it is possible to relay full screen images of a demonstration set-up, a procedure, or an experiment. LFV may be used also to transfer the image of a conventional white-board.

The complete system can be operated by the teacher without the assistance of an additional operator or technician. However, at the remote site, a system manager should be responsible for the equipment and the correct course of the session. The latter function may, however, be delegated to one of the students. Further it is worth noting that the system is suitable for candid-classroom operation. This involves simultaneous teaching of a live class and one or more remote classes. In the live class a large screen projection of the shared area must be provided.

Table 8.4 shows how the features of the tele-education system correlate with the features of conventional face-to-face teaching. Additionally, remember that the system is characterized by short transfer times of visuals and is easy to operate. Compared to face-to-face teaching the system offers enhanced facilities like filing, editing, re-use, and hardcopying of materials. If desired the lecture can be video-taped through an additional VGA-to-PAL conversion device. A limitation is that the frame-rates of LFV are significantly lower than real-time video. For instance at 9600 bit/s the rates range from 0.2 to 1 Hz, depending on the image size.

At extremely low bitrates (1200 bit/s) the system can still be effectively utilized, provided that prepared presentations are used to exploit the advantages of the pretransfer mechanism. At this bitrate the transfer of on-line inputted text, graphics, or handwriting is still constrained by the human handling speed.
Table 8.4 Tele-education system facilities compared to conventional face-to-face teaching features

<table>
<thead>
<tr>
<th>Teaching feature</th>
<th>Tele-education correlate</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>classroom: 'presence' and participants</td>
<td>LFV; audio</td>
<td>two-way LFV: small, medium, large screen; two-way audio</td>
</tr>
<tr>
<td>voice communication</td>
<td>audio</td>
<td></td>
</tr>
<tr>
<td>non-verbal communication</td>
<td>LFV</td>
<td></td>
</tr>
<tr>
<td>projector (slide or overhead)</td>
<td>still video pictures by document camera stand</td>
<td>256 colours or bi-level raster-scanned *)</td>
</tr>
<tr>
<td>writing &amp; drawing</td>
<td>writing tablet; text &amp; graphics (keyb.&amp; mouse)</td>
<td>*)</td>
</tr>
<tr>
<td>presentation of visuals</td>
<td>stored sequence of compound pages</td>
<td>non-sequential order allowed *)</td>
</tr>
<tr>
<td>demonstration (object, procedure, experiment)</td>
<td>LFV full screen</td>
<td></td>
</tr>
<tr>
<td>writing on visuals</td>
<td>on-line input of text, graphics, handwriting</td>
<td>*)</td>
</tr>
<tr>
<td>pointing on visuals</td>
<td>movable pointer</td>
<td>*)</td>
</tr>
<tr>
<td>improvised explanation</td>
<td>scratch window</td>
<td>*)</td>
</tr>
<tr>
<td>interaction</td>
<td>interrupt request; primary/secondary mode</td>
<td>primary dictates floor passing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*) input to shared area</td>
</tr>
</tbody>
</table>

8.8 A class-room experiment

Using the initial version of the above described system, a class-room experiment was performed at the University of the Netherlands Antilles (UNA) at Curacao. A group of eight undergraduate students, present in a lecture-room, was being taught from an office at another part of the campus. These lectures (as it happened, on advanced telecommunications) were the last ones of a regular semester course. Hence it was of particular
Fig. 8.11 Overview of the classroom situation during the experiment described

importance to these students to attend the sessions, as there were no additional lectures scheduled before the corresponding preliminary examination. This guaranteed a serious and highly motivated participation in this educational experiment. If it would appear that the tele-education sessions were not effective enough, the contingency to repeat the lectures by face-to-face sessions was foreseen. This was not necessary, however. Each of the lectures was followed by an evaluation involving teacher and students to complete evaluation forms individually. Its purpose was to investigate the usability of the system for real-time distance teaching and evaluate technical and user aspects.

Fig. 8.11 gives an overview of the situation in the lecture room. The configuration of the teacher’s terminal was as depicted in Fig. 8.1. The terminals were connected to each other by dedicated lines for audio and data. The latter comprised a simple ‘null-modem’ connection (full duplex, 9600 bit/s). Only low-resolution LFV with a frame rate of about 0.1 Hz was possible with this initial prototype. No form of pretransfer had been implemented. Before the start of the session only a short user instruction had to be given, since the participating students were familiar with PC applications. Some of them in fact contributed to the development of the prototype.

The following conclusions and recommendations summarize the major evaluation results. The system appears to be appropriate to conduct distance lectures at a professional level. The multimedia concept used enables a way of teaching close to conventional
face-to-face methods, although a few improvements may be necessary to increase the effectiveness. No major problems were reported concerning transfer times of pages or objects. From the evaluation the following recommendations resulted:

1) The dimensions of the classroom display should be enlarged ($\geq 1$ m).
2) Pictures and LFV images should be displayed in colour.
3) The LFV image size should be enlarged for two reasons: i) to enable a more detailed inspection of the class situation by the teacher; ii) to be able to transfer the video image of a whiteboard, object, demo set-up, etc. to the students.
4) The frame-rate of the LFV image should be increased.
5) The scrolling function should be improved. Scrolling in smaller steps (e.g., of about one text line) probably would give better results than the current large steps (of one half of the screen area).
6) Text line spacing should be adjustable.
7) A hardcopy facility was requested, to deliver a printout to the students of the presented pages including on-line added annotations and supplements.
8) Concerning the user interface more flexibility regarding possibilities to switch between the deeper layers of the menu structure should be added.

Some of these recommendations have been implemented in the developed system described in this chapter. The LFV image can be enlarged to (nearly) full screen dimensions. The LFV frame-rate is significantly increased due to the development of an efficient coding scheme (see Chapter 4).

### 8.9 Summary and conclusions

Real-time tele-education based on multimedia concepts permits realization of more effective distance courses. Application in third-world countries usually enforces a number of specific conditions and constraints, related to the available telecommunication infrastructure, the local level of serviceability regarding complex equipment, and the constraints imposed by the educational institute involved. Study of previous work related to real-time tele-education revealed that little research has been spent on the development of low-cost real-time systems which can be used in combination with narrowband telecommunication links.
This chapter described such a system, suitable for point-to-point and for multipoint configurations. Structured presentations consisting of compound pages may be presented in combination with on-line input and editing. The system applies the following media: typographic text, structured graphics, raster-scanned pictures, handwriting, audio, and low frame-rate video (LFV). To support the production and management of educational materials an extensive and powerful set of editing, macro, and file management functions is provided. The shared areas of all screens are kept synchronized throughout the session, while transfer times are minimized due to the application of pretransfer methods. A multipoint floor control mechanism supports the interaction process and regulates audio, and video distribution as well as shared-screen access permission.

A classroom experiment has been performed at the university premises at Curaçao. Evaluation results revealed that the system is appropriate for the conduction of distance lectures at a professional level. The multimedia concept used enables a way of teaching close to conventional face-to-face methods.
References in Chapter 8


Arnback, J.C. [88], "Trans-border communications networks: technology-driven or social innovation?", European Management Journal, Vol.6, No.4, pp.330-337

Bates, A.W. [88], "Technology for distance education: a 10-year prospective", Open Learning, pp.3-12, November 1988


Judice, C.N., D. LeGall [87], "Telematic services: are we ready?", IEEE Comm. Mag., Vol.25, pp.19-29, 1987

Kegel, A., J.H. Bons [77], "On the digital processing and transmission of handwriting and sketching", Proc. EUROCON'77, Venice, 1977, pp.2.4.6.1-6
Lalor, G.C., C.Murrett [86], *Report on the University of the West Indies Distance Teaching experiment*, Kingston: University of the West Indies, 1986


Rosenthal, L. [87], *The Eastern Caribbean Center Telecommunications Project*, University of the Virgin Islands, February 1987


Tsuruta, S., K.Matsuhashi, M.Mera, K.Niwa [86], "Intelligent communication terminal for integrating voice, data and video signals", *Proc. ICC’86*, pp.1509-1513, 1986


In this thesis we addressed the possibilities to solve a frequently occurring problem in education by employing modern telematic technologies. In many countries there are situations in which groups of people lack any form of higher education due to, e.g., geographical barriers. The conventional approach to this problem is the correspondence course, which suffers a number of typical disadvantages. As a consequence, in general this type of distance education is less effective than normal face-to-face courses. Moreover, regular educational institutes are not adequately equipped to extend their facilities with such a form of distance education. However, real-time tele-education using computer technology in combination with telecommunications enables a way of teaching which approximates the normal face-to-face situation rather closely. Hence this class of systems can be integrated more easily into the educational organizations of existing institutes. Besides, the high level of interactivity without the disadvantage of a long turn-around time contributes to higher learning effectiveness.

The design of a tele-education system has been described, particularly suitable for situations similar to that of the Netherlands Antilles. The archipelago of the Netherlands Antilles consists of a number of small islands with mutual distances of up to 800 kilometres. Only the island of Curacao accommodates an institute of higher education: the University of the Netherlands Antilles (UNA). From an analysis of the specific requirements, limiting conditions, and constraints, a PC-based system has been developed, with the maximum functionality implemented in software. This results in a relatively low-cost solution and requires only a minimal level of serviceability. The available telecommunication bandwidth does not allow the application of real-time video. Therefore the adopted approach was to develop a low frame-rate video method. In addition, the system applies electronic variants of the regular teaching media, using a writing tablet, a
document camera, an audio channel, computer input, and several extra features. Educational material for the lecture can be produced in advance, for which purpose extensive editing and file management functions are incorporated. It is possible to teach various remote groups simultaneously in a multipoint configuration. In real-time tele-education it is of vital importance that end-to-end transfer times of visuals are as small as possible. A method has been developed to reduce these delays to virtually zero, even in the case of very small bandwidths. An initial version of the system designed has been involved in a classroom experiment at the UNA. From evaluation results it may be concluded that the multimedia concept developed meets the requirements and it is a viable approach to tele-education in the specific situations aimed at.

9.1 Major Results and Conclusions

1) In Chapter 1 we gave a classification of tele-education methods, involving two main categories: mail-based and real-time tele-education. Mail-based methods employ computer-mediated teaching based on electronic mail services. Compared to the conventional correspondence course, the response times are reduced significantly and mutual student interaction is made possible. However, immediate feedback is only possible by real-time methods.

2) Chapter 2 presented an overview of the various tele-education methods implemented in real educational situations, including a number of prominent examples. A description was given of one-way systems, two-way audio, two-way video, two-way graphics & text and two-way audio, slow-scan video and two-way audio, electronic class and virtual class.

3) In addition, Chapter 2 stressed the importance of a systematic and problem-oriented approach. In this context the significance of educational technology has been emphasized. This involves a systems approach consisting of the following general phases: problem definition, problem analysis and selection of a solution, development, implementation and testing, evaluation and revision. Furthermore, some general considerations and guidelines were presented important to the application of a tele-education solution.
4) Chapter 3 studied specific situations where real-time tele-education is an effective solution. The University of the Netherlands Antilles has been used as a case representing a broader range of similar situations, in order to analyze problem areas and to study various approaches to a solution. It has been concluded that real-time tele-education offers the best prospects to realize effective and successful educational services aimed at remote communities. This method is the closest replica of regular face-to-face teaching. Consequently only minor adoptions of teaching methods and curricula are required, yielding the advantages of relatively low operational costs and a rather low psychological threshold to get such a method introduced. Besides two-way audio communication a visual component is imperative in real-time tele-education. Where the lack of wideband telecommunication channels prevents the application of Instructional Television (ITV), the use of narrowband visual signals should be applied. Therefore a requirement analysis has been performed, aimed at the development of a multimedia PC-based system including audio, graphics and reduced-video services. Furthermore the impact of the local telecommunication infrastructure on such a system has been assessed and two main network concepts have been identified: a terrestrial multipoint network employing the capacities of the Public Switched Telephone Network (PSTN) and a satellite-based concept implying the lecture broadcast to Receive-Only remote stations with the possibility of phone-in feedback via the PSTN. Besides, a conceptual system architecture has been discussed and a number of critical technologies identified.

5) In Chapter 4 a Low Frame-rate Video Coding method (LFVC) was presented, which has been developed as a sub-system for a real-time tele-education terminal. It encompasses a fast and efficient coding scheme, appropriate for software implementation, and aimed at narrow bandwidths. The coding scheme is useful for use in a multimedia tele-education terminal, when there is a requirement for implementation on standard microcomputer equipment without dedicated hardware. Such a subsystem allows the teacher to watch the classroom (and vice versa) during the session, by near full-screen image sequences. An analysis on the relations between coder complexity, compression factor, CPU-load and transfer time has been performed, revealing the need to develop a coding scheme with significant lower complexity than standard methods like JPEG DCT. For this purpose the FEVQ (Fast and Efficient Vector Quantization) algorithm was developed, applying LUT-based MSVQ on 2x2 pixel blocks in combination with DPCM on the block means and Modified Tunstall Coding. Low complexity interframe compression can be achieved by conditional replenishment. A two-level approach was adopted: replenishment of
2x2 pixel blocks and of 128x96 pixel blocks (called sub-images). With sub-image replenishment based on a sample check mechanism, a very efficient encoder implementation was shown to be feasible.

To verify the results obtained, a complete software-based codec has been realized, achieving acceptable image end-to-end transfer times even at narrow bandwidths (e.g., 9600 bit/s).

6) High-speed Huffman coding, implemented in software, requires large amounts of memory. In Chapter 5, it has been shown that variable-to-fixed-length coding schemes like Tunstall coding outperform Huffman coding not only with respect to the coding rate in the case of a low entropy source, but also with respect to complexity. Moreover, variable-to-fixed-length coding schemes are less sensitive to transmission errors. A modified version of the Tunstall coding method has been proposed (MTC), with a coding rate closer to the source entropy due to a faster convergence rate. In conclusion, MTC is an attractive choice for software-implemented entropy coding in the event of limited alphabet sources with a skew probability distribution.

7) A number of practical improvements of present telewriting systems were proposed with respect to tele-education applications. It has been shown in Chapter 6 that Differential Chain Coding (DCC), a technique used to encode telewriting, can be enhanced with a fill facility in a way that graphical editing functions remain feasible. The extension or editing of on-screen-displayed objects by means of an electronic writing tablet causes an eye-to-hand coordination problem. A Relative Coordinate System (RCS) facility has been proposed by which the user is able to position handwritten graphics on screen at a precisely defined location.

8) Further DCC was shown to be a particular, but sub-optimal form of Modified Tunstall Coding. It is possible to optimize DCC by applying an extended set of codewords, representing segments of relative or absolute vectors. It has been shown that as a result, there is a significant reduction of the coding rate. This efficient method preserves the typical robustness of DCC against transmission errors.

9) To reduce end-to-end delays in multimedia communications, Chapter 7 described a time-shifting method, referred to as pretransfer, which exploits periods of low channel utilisation during the session for automatic transmission of page sequences in background. Point-to-point and multipoint application protocols have been dis-
cussed, resulting in virtual zero transfer times. The pretransfer concept may also be applied in multicast situations, for which various page scheduling methods have been analyzed.

10) Finally, Chapter 8 described the design and implementation of a PC-based real-time tele-education system aimed at narrow bandwidths. The system was developed taking into account the specific requirements and conditions imposed by the typical situation of a small-scale university in a third-world country. The technical developments as proposed in the Chapters 4-7 have been incorporated in the design. From the results of a classroom experiment performed at the University of the Netherlands Antilles at Curaçao it may be concluded that the multimedia concept proposed is appropriate for the conduction of distance lectures at a professional level.

9.2 Recommendations on Future Work

Some recommendations on future work related to the real-time tele-education technology developed are suggested below:

1) The developed tele-education system should be more extensively tested by performing a pilot experiment. Such a small-scale trial involves the realization of a few tele-courses during (a part of) one semester. Preferably this should be done in a multipoint configuration in combination with a live class ('candid classroom'). The objective is to evaluate user aspects and technical system performance. In addition, such a test will be useful to gain experience with respect to financial budgets, organizational aspects, resulting learning effectiveness, and training effort required. Further, issues like the impact of power outages, link dropouts, serviceability, etc., might be studied. Such a pilot experiment would not only provide data in order to optimize the tele-education system, but yield useful information necessary for the implementation of a larger scale service.

2) It is recommended to investigate whether the following media should be supported as well:
   a) Electronic whiteboard. This medium enables the use of a whiteboard in a tele-education situation with the advantages of improved resolution and reduced transfer
times, compared to low frame-rate video sequences of a normal whiteboard. Moreover, it is expected that most teachers will prefer this medium to the electronic writing tablet for on-line handwriting and linedrawing.

b) Remote-controlled observation camera. To observe the classroom situation the teacher should be able to control the video camera situated in the classroom (zooming and scrolling functions).

c) Captured PC-screens. It should be considered to add captured PC-screens to the visuals which can be communicated. This is particularly useful for courses in computer science and computer systems, but all kinds of computer-generated screens may be transmitted in this way. Capturing utilities are available which immediately convert the screens in a compressed format.

3) As has been discussed in Sections 3.3.1 and 8.6.1, an interesting option would be to develop a multicast variant. In this every remote group would receive the same continuous signal (consisting of audio and data) originating from the primary station. A viable approach to interactivity is PSTN phone-in feedback by scribophone technology. Such a configuration might be attractive for application in regions in Western Europe as well. For instance, datacasting piggybacked on a D2MAC satellite video channel is a way to reach a large and widely dispersed target group, with efficient utilisation of telecommunication facilities. Pretransfer methods can be applied to minimize end-to-end transfer times.
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADPCM</td>
<td>Adaptive Pulse Code Modulation</td>
<td>HDTV</td>
<td>High Definition Television</td>
</tr>
<tr>
<td>AGC</td>
<td>Automatic Gain Control</td>
<td>HVS</td>
<td>Human Visual System</td>
</tr>
<tr>
<td>ALU</td>
<td>Application Layer User</td>
<td>ISDN</td>
<td>Integrated Services Digital Network</td>
</tr>
<tr>
<td>AMBTC</td>
<td>Absolute Moment Block Truncation Coding</td>
<td>ISO</td>
<td>International Standardization</td>
</tr>
<tr>
<td>AMTC</td>
<td>Adaptive Modified Tunstall Coding</td>
<td>ITV</td>
<td>Instructional Television</td>
</tr>
<tr>
<td>ARQ</td>
<td>Automatic Repeat Request</td>
<td>JPEG</td>
<td>Joint Photographers’ Expert Group</td>
</tr>
<tr>
<td>ATM</td>
<td>Asynchronous Transfer Mode</td>
<td>KLT</td>
<td>Karkhonen Loeve Transform</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
<td>LAN</td>
<td>Local Area Network</td>
</tr>
<tr>
<td>B-ISDN</td>
<td>Broadband ISDN</td>
<td>LFV</td>
<td>Low Frame-rate Video</td>
</tr>
<tr>
<td>BTC</td>
<td>Block Truncation Coding</td>
<td>LFVC</td>
<td>Low Frame-rate Video Coding</td>
</tr>
<tr>
<td>CCITT</td>
<td>Comité Consultatif International Télégraphique et Téléphonique</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD-I</td>
<td>Compact Disk Interactive</td>
<td>LUT</td>
<td>Look-Up Table</td>
</tr>
<tr>
<td>CD-ROM</td>
<td>Compact Disk Read-Only Memory</td>
<td>MHEG</td>
<td>Multimedia and Hypermedia</td>
</tr>
<tr>
<td>CNPPS</td>
<td>Current &amp; Next Page Preemptive Scheduling</td>
<td>MPEG</td>
<td>Motion Pictures Expert Group</td>
</tr>
<tr>
<td>CPPS</td>
<td>Current Page Preemptive Scheduling</td>
<td>MSE</td>
<td>Mean Squared Error</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
<td>MSVQ</td>
<td>Mean Separating Vector Quantization</td>
</tr>
<tr>
<td>CR</td>
<td>Conditional Replenishment</td>
<td>NPPS</td>
<td>Next Page Preemptive Scheduling</td>
</tr>
<tr>
<td>CSCW</td>
<td>Computer Supported Cooperative Work</td>
<td>NPS</td>
<td>Non-Preemptive Scheduling</td>
</tr>
<tr>
<td>DMI</td>
<td>Direct Manipulation Interface</td>
<td>MTC</td>
<td>Modified Tunstall Coding</td>
</tr>
<tr>
<td>DMS</td>
<td>Discrete Memory Source</td>
<td>NTSC</td>
<td>National Television Standard</td>
</tr>
<tr>
<td>DPCM</td>
<td>Differential Pulse Code</td>
<td>ODA</td>
<td>Office Document Architecture</td>
</tr>
<tr>
<td>DMAC</td>
<td>Digital + Analog Multiplexed Components</td>
<td>OSI</td>
<td>Open Systems Interconnection</td>
</tr>
<tr>
<td>DMI</td>
<td>Direct Manipulation Interface</td>
<td>PAL</td>
<td>Phase Alternating Line (TV standard)</td>
</tr>
<tr>
<td>DMS</td>
<td>Discrete Memory Source</td>
<td>PBX</td>
<td>Private Branch Exchange</td>
</tr>
<tr>
<td>DPCM</td>
<td>Differential Pulse Code</td>
<td>PCM</td>
<td>Pulse Code Modulation</td>
</tr>
<tr>
<td>DMAC</td>
<td>Digital + Analog Multiplexed Components</td>
<td>PRP</td>
<td>Pretransfer Process</td>
</tr>
<tr>
<td>DOS</td>
<td>Disk Operating System</td>
<td>PTM</td>
<td>Pretransfer Machine</td>
</tr>
<tr>
<td>DVI</td>
<td>Digital Video Interactive</td>
<td>PTS</td>
<td>Pretransfer Supervisor</td>
</tr>
<tr>
<td>DWHT</td>
<td>Discrete Walsh Hadamard Transform</td>
<td>PSNR</td>
<td>Peak Signal to Noise Ratio</td>
</tr>
<tr>
<td>EGA</td>
<td>Enhanced Graphics Adapter</td>
<td>PSTN</td>
<td>Public Switched Telephone Network</td>
</tr>
<tr>
<td>FDDI</td>
<td>Fiber Distributed Data Interchange</td>
<td>RCS</td>
<td>Relative Coordinate System</td>
</tr>
<tr>
<td>FEC</td>
<td>Forward Error Correction</td>
<td>RGB</td>
<td>Read, Green, Blue</td>
</tr>
<tr>
<td>FEVQ</td>
<td>Fast &amp; Efficient Vector Quantization</td>
<td>SBC</td>
<td>Subband Coding</td>
</tr>
<tr>
<td>FM</td>
<td>Frequency Modulation</td>
<td>SCPC</td>
<td>Single Channel Per Carrier</td>
</tr>
<tr>
<td>FV</td>
<td>Fixed-to-Variable-length coding</td>
<td>SGML</td>
<td>Standard Generalized Markup Language</td>
</tr>
<tr>
<td>HCI</td>
<td>Human Computer Interface</td>
<td>SNR</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>HDMAC</td>
<td>High Definition Multiplexed Analog Components</td>
<td>SSB</td>
<td>Shared Screen Buffer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SSTV</td>
<td>Slow Scan Television</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TC</td>
<td>Tunstall Coding</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------</td>
<td>---------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>TVRO</td>
<td>Television Receive Only</td>
<td>VF</td>
<td>Variable-to-Fixed-length coding</td>
</tr>
<tr>
<td>UA</td>
<td>University of Aruba</td>
<td>VGA</td>
<td>Video Graphics Adapter</td>
</tr>
<tr>
<td>UHB</td>
<td>User Home Base</td>
<td>VLSI</td>
<td>Very Large Scale Integration</td>
</tr>
<tr>
<td>UHF</td>
<td>Ultra High Frequency</td>
<td>VPIC</td>
<td>Visual Pattern Image Coding</td>
</tr>
<tr>
<td>UI</td>
<td>User Interface</td>
<td>VQ</td>
<td>Vector Quantization</td>
</tr>
<tr>
<td>UNA</td>
<td>University of the Netherlands</td>
<td>VSAT</td>
<td>Very Small Aperture Terminal</td>
</tr>
<tr>
<td></td>
<td>Antilles</td>
<td>VZT</td>
<td>Virtual Zero Transfer Time</td>
</tr>
<tr>
<td>VCR</td>
<td>Video Cassette Recorder</td>
<td>WAN</td>
<td>Wide Area Network</td>
</tr>
<tr>
<td>VDU</td>
<td>Visual Display Unit</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A Software-based implementation of Extended Hamming code

The (128,120) Extended Hamming code\(^1\) enables the correction of one bit error or the detection of two bit errors in a 120 bits data block. Eight error check bits need to be transmitted in addition to the 120 data bits, which can be determined employing the \(h\) matrix:

\[
\begin{bmatrix}
    h_{11} & h_{21} & h_{31} & \ldots & h_{k1} \\
    h_{12} & h_{22} & h_{32} & \ldots & h_{k2} \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    h_{17} & h_{27} & h_{37} & \ldots & h_{k7}
\end{bmatrix}
\]

(1)

Here \(k = 120\), and \(h_{ij} \in \{0,1\}\). The columns \([ h_{i1} \ h_{i2} \ h_{i3} \ \ldots \ h_{i7} ]\) of \(h\) represent the binary representations of all integers in \(\{1,\ldots,127\}\) except for powers of 2, in ascending order. Let the 120 data bits be denoted by a vector \(D = [d_1 \ d_2 \ \ldots \ d_{120}]\). Then 7 of the 8 error check bits are generated according to:

\[
C = [c_1 \ c_2 \ \ldots \ c_7] = Dh^T
\]

(2)

The remaining check bit is the parity calculated over the 120 data bits and the 7 error check bits. Let \(R = [r_1 \ r_2 \ \ldots \ r_{120} \ \ldots \ r_{127}]\) represent the vector of the received bits, excluding the parity bit. Then at the receiver the syndrome \(S\) is calculated by:

\[
S = HR^T
\]

(3)

\(H\) is the matrix composed of \(h\) and the unity matrix, according to:

---

\[ H = \begin{bmatrix}
 h_{11} & h_{12} & h_{13} & \cdots & h_{1k} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 h_{21} & h_{22} & h_{23} & \cdots & h_{2k} & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 h_{71} & h_{72} & h_{73} & \cdots & h_{7k} & 0 & 0 & 0 & 0 & 0 & 0 & 1 
\end{bmatrix} \]

(4)

If \( S^T = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0] \), the number of bit errors is equal to 0. If \( S^T \neq [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0] \), one or more errors occurred. Then \( S \) equals one of the columns of \( H \), with column index \( j \), \( j \in \{1, \ldots, 127\} \). In the event of one bit error, the position of the faulty bit is equal to \( j \). Hence this error can be corrected by the receiver. If two bit errors occur (following from \( S^T \neq [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0] \) and a correct parity bit), no correction is possible, leading to the rejection of the received data block.

To implement this error correction mechanism in software in a computationally efficient way, the following method can be applied. Blocks of 120 data bits are divided each into fifteen consecutive bytes, denoted by vectors \( d_m \) of dimension 8, with \( m \in \{0, \ldots, 14\} \). Let \( h_m \) represent a sub-matrix of \( h \) according to:

\[ h_m = \begin{bmatrix}
 h_{(8m+1)1} & h_{(8m+1)2} & h_{(8m+1)3} & \cdots & h_{(8m+1)7} \\
 h_{(8m+2)1} & h_{(8m+2)2} & h_{(8m+2)3} & \cdots & h_{(8m+2)7} \\
 \vdots & \vdots & \vdots & \cdots & \vdots \\
 h_{(8m+7)1} & h_{(8m+7)2} & h_{(8m+7)3} & \cdots & h_{(8m+7)7} 
\end{bmatrix} \]

(5)

Then operation (2) can be rewritten as follows:

\[ C = d_0 h_0^T \oplus d_1 h_1^T \oplus \cdots \oplus d_{14} h_{14}^T \]

(6)

The fifteen terms of (6) can be efficiently calculated by look-up tables using \( d_m \) as entry. \( C \) results from modulo 2 accumulation of the fifteen 7-bits return values. It is easily verified that the calculation of the parity bit can be integrated in these tables by the adoption of 8-bit wide table values. At the receiver \( S \) may be determined by a similar approach. In this case the return value should be 0 (no errors), \( j \) (the position of the faulty bit in the case of one error), or \( > 127 \) (more than 2 errors).

In conclusion, this LUT-based FEC approach requires only \( \frac{16}{120} \) operations per databit.
B List of menu command options

The following table lists all menu command options including icons and indicators. The codes in the first column indicate the position of the item in the menu hierarchy. The main menu (1..8) enables the selection of one of eight sub-menus. For example, choosing 'screen functions' (number 1) pops up the sub-menu with the options 1.1 to 1.7. Subsequently, selection of 'zoom & scroll functions' (number 1.2) results in the options 1.2.1 to 1.2.7, etc. A menu bar maximally contains 9 options. The second column lists the commands, sub-menus (printed in capitals), icons, and indicators. Icons are used to identify the current menu level. Indicators are small icon-like images presenting the current value of a parameter (may be a boolean, a colour, a scalar, a width, etc.). Indicators and icons are both shown in italics.

<table>
<thead>
<tr>
<th>Menu tree command</th>
<th>Command / icon / indicator</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SCREEN FUNCTIONS</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>COMMUNICATIONS</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>FILE SERVICES</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>TEXT</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>GRAPHICS</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>PICTURES</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>HANDWRITING</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>MACROS</td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>screen</td>
<td>Toggles scratch mode</td>
</tr>
<tr>
<td>1.2</td>
<td>ZOOM &amp; SCROLL</td>
<td>Items entered in scratch mode are not saved and the current screen is not altered. This mode can be left by the update command as well.</td>
</tr>
<tr>
<td>1.3</td>
<td>scratch</td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>VIDEO</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>MARKERS</td>
<td></td>
</tr>
<tr>
<td>Menu tree command</td>
<td>Command / icon / indicator</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>1.6</td>
<td>update</td>
<td>Screen update</td>
</tr>
<tr>
<td>1.7</td>
<td>EDIT</td>
<td>To remove footprints caused by editing</td>
</tr>
<tr>
<td>1.2.1</td>
<td>zoom &amp; scroll</td>
<td>Editing all types of objects</td>
</tr>
<tr>
<td>1.2.2</td>
<td>in</td>
<td>Zoom in (2 levels)</td>
</tr>
<tr>
<td>1.2.3</td>
<td>out</td>
<td>Zoom out (2 levels)</td>
</tr>
<tr>
<td>1.2.4</td>
<td>left</td>
<td>Scroll to the left</td>
</tr>
<tr>
<td>1.2.5</td>
<td>right</td>
<td>Scroll to the right</td>
</tr>
<tr>
<td>1.2.6</td>
<td>up</td>
<td>Scroll up</td>
</tr>
<tr>
<td>1.2.7</td>
<td>down</td>
<td>Scroll down</td>
</tr>
<tr>
<td>1.4.1</td>
<td>video</td>
<td>Switch to large-size video</td>
</tr>
<tr>
<td>1.4.2</td>
<td>large</td>
<td>Switch to medium-size video</td>
</tr>
<tr>
<td>1.4.3</td>
<td>medium</td>
<td>Switch to small-size video</td>
</tr>
<tr>
<td>1.4.4</td>
<td>small</td>
<td>Suppress video</td>
</tr>
<tr>
<td>1.4.5</td>
<td>suppress</td>
<td>Toggle between high and medium resolution</td>
</tr>
<tr>
<td>1.4.6</td>
<td>resolution</td>
<td>Adjust brightness</td>
</tr>
<tr>
<td>1.4.7</td>
<td>brightness</td>
<td>Brightness indicator</td>
</tr>
<tr>
<td>1.4.8</td>
<td>contrast</td>
<td>Adjust contrast</td>
</tr>
<tr>
<td>1.4.9</td>
<td>contrast</td>
<td>Contrast indicator</td>
</tr>
<tr>
<td>1.5.1</td>
<td>markers</td>
<td>(De)activate left pointer</td>
</tr>
<tr>
<td>1.5.2</td>
<td>&lt;--</td>
<td>(De)activate right pointer</td>
</tr>
<tr>
<td>1.5.3</td>
<td>--&gt;</td>
<td>Clicking the right mouse button (or Enter key) switches between left and right pointer.</td>
</tr>
<tr>
<td>1.5.4</td>
<td>marker</td>
<td>Marker function</td>
</tr>
<tr>
<td>1.5.5</td>
<td>color</td>
<td>The centre mouse button (or Ins key) toggles between large and small marking block-size. The right mouse button (or Enter key) activates actual marking.</td>
</tr>
<tr>
<td>1.7.1</td>
<td>edit</td>
<td>Define marking color</td>
</tr>
<tr>
<td>1.7.2</td>
<td>move</td>
<td>Move selected objects</td>
</tr>
<tr>
<td>1.7.3</td>
<td>copy</td>
<td>Copy selected objects</td>
</tr>
<tr>
<td>Menu tree command</td>
<td>Command / icon / indicator</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>1.7.4</td>
<td>erase</td>
<td>Erase selected objects</td>
</tr>
<tr>
<td>1.7.5</td>
<td>backgr. color</td>
<td>Define background colour</td>
</tr>
<tr>
<td>1.7.6</td>
<td>pause</td>
<td>Pausemark input/edit mode. A pausemark can be inserted by the INS key after the highlighted object. Object highlighting can be rotated using the arrow keys. 'Highlighting' of a pausemark is displayed by an indicator at the lower righthand corner. A pausemark can be deleted by the DEL key.</td>
</tr>
<tr>
<td>2.1</td>
<td><em>com</em></td>
<td>Activate audio-visual attention signal at other station</td>
</tr>
<tr>
<td>2.2</td>
<td>attend</td>
<td>Reset attention signal. Used for resetting of incoming signal or stopping of outgoing signal.</td>
</tr>
<tr>
<td>2.3</td>
<td>attres</td>
<td>Select one or more secondary stations. Only in primary mode. Used in combination with call/disconnect, attention, and floor allocation.</td>
</tr>
<tr>
<td>2.4</td>
<td>select</td>
<td>The floor is given to the selected secondary station (interaction mode). Selection may be the result of a floor request from the remote station by its attention command. Subsequent activation of this command deselects the station and transfers the floor to the next selection (chronologically). If no selections are left, the system switches back to discourse mode.</td>
</tr>
<tr>
<td>2.5</td>
<td>floor</td>
<td>Reset selection and switch to discourse mode</td>
</tr>
<tr>
<td>2.6</td>
<td>reset sel</td>
<td>Link management functions</td>
</tr>
<tr>
<td>2.7</td>
<td>LINK</td>
<td>Indicator array for all connected stations. Presents calling progress (calling, link established, disconnected), and selection.</td>
</tr>
<tr>
<td>2.8</td>
<td>remote stations</td>
<td>Down counter indicating number of seconds</td>
</tr>
<tr>
<td>Menu tree command</td>
<td>Command / icon / indicator</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>2.7.1</td>
<td><em>com</em></td>
<td>to go until the sent object or page is displayed at the remote station. Note that this feature is not necessary in the case of pretransfer.</td>
</tr>
<tr>
<td>2.7.2</td>
<td>link</td>
<td>Call/disconnect communication link. In primary mode this command may involve multipoint call set-up (corresponding to the current selection of remote stations)</td>
</tr>
<tr>
<td>2.7.3</td>
<td>pretransfer</td>
<td>Enable/disable pretransfer</td>
</tr>
<tr>
<td>2.7.4</td>
<td>mon</td>
<td>Communication monitor window</td>
</tr>
<tr>
<td>2.7.5</td>
<td>primary</td>
<td>Selects primary status after entering a valid user-name/password combination.</td>
</tr>
<tr>
<td>2.7.6</td>
<td>scalign</td>
<td>Shared-screen alignment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Upline all current screen parameters and objects of communicating stations</td>
</tr>
<tr>
<td>3.1</td>
<td><em>pres: name</em></td>
<td>Current presentation name</td>
</tr>
<tr>
<td>3.2</td>
<td>SAVE</td>
<td>Save page functions</td>
</tr>
<tr>
<td>3.3</td>
<td>get</td>
<td>Load page of current presentation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asks for the page index in the current presentation sequence</td>
</tr>
<tr>
<td>3.4</td>
<td>next</td>
<td>Load next page of current presentation</td>
</tr>
<tr>
<td>3.5</td>
<td>previous</td>
<td>Load previous page of current presentation</td>
</tr>
<tr>
<td>3.6</td>
<td>PRESENTATION</td>
<td>Presentation functions</td>
</tr>
<tr>
<td>3.7</td>
<td>overlay</td>
<td>Overlay a page</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asks for a presentation name and a page index</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Remarks: The first page of a presentation has index 0. The directory command (Presentation submenu) gives an overview of the pages in the current presentation.</td>
</tr>
<tr>
<td>3.2.1</td>
<td><em>pres: name</em></td>
<td>Current presentation name</td>
</tr>
<tr>
<td>3.2.2</td>
<td>next</td>
<td>Save page as the next in current presentation</td>
</tr>
<tr>
<td>3.2.3</td>
<td>replace</td>
<td>Replace a page of the presentation by the</td>
</tr>
<tr>
<td>Menu tree command</td>
<td>Command / icon / indicator</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>3.2.4</td>
<td>insert</td>
<td>current page. Insert current page in the current presentation. Asks for a page index. The current page gets the entered page index and all subsequent pages get an incremented index.</td>
</tr>
<tr>
<td>3.2.5</td>
<td>message</td>
<td>Enter a page message. To be used for identification purposes. The message is displayed in the presentation directory list.</td>
</tr>
<tr>
<td>3.2.6</td>
<td>label</td>
<td>Toggles page label</td>
</tr>
<tr>
<td>3.6.1</td>
<td>pres: name</td>
<td>Current presentation name</td>
</tr>
<tr>
<td>3.6.2</td>
<td>change</td>
<td>Change presentation name</td>
</tr>
<tr>
<td>3.6.3</td>
<td>delete</td>
<td>Delete a page from the current presentation. Asks for a page index. All subsequent pages get a decreased (-1) index</td>
</tr>
<tr>
<td>3.6.4</td>
<td>pdir</td>
<td>Display directory of the current presentation. Displays page index and page message of all pages</td>
</tr>
<tr>
<td>3.6.5</td>
<td>carsl</td>
<td>Display the pages of current presentation cyclically. Only for demo purposes.</td>
</tr>
<tr>
<td>3.6.6</td>
<td>chdir</td>
<td>Change DOS directory</td>
</tr>
<tr>
<td>3.6.7</td>
<td>ddir</td>
<td>Display DOS directory</td>
</tr>
<tr>
<td>4.1</td>
<td>text</td>
<td>Indicator for current text colour</td>
</tr>
<tr>
<td>4.3</td>
<td>EDIT</td>
<td>Text edit functions</td>
</tr>
<tr>
<td>4.4</td>
<td>font</td>
<td>Change font.</td>
</tr>
<tr>
<td>4.5</td>
<td>size</td>
<td>Change text size. Asks for a size number</td>
</tr>
<tr>
<td>Menu tree command</td>
<td>Command / icon / indicator</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>4.6</td>
<td>color</td>
<td>(0..2). The available sizes 0,1,2 correspond with 20, 10 and 5 lines per screen, respectively.</td>
</tr>
<tr>
<td>4.7</td>
<td>clear</td>
<td>Change text colour</td>
</tr>
<tr>
<td>4.3.1</td>
<td>text</td>
<td>Clear all text on screen</td>
</tr>
<tr>
<td>4.3.2</td>
<td>mode</td>
<td>Indicator for current edit mode.</td>
</tr>
<tr>
<td>4.3.3</td>
<td>mode</td>
<td>Edit mode: object or field. In field mode the user may select a group of objects inside a specifiable rectangle. In object mode the selection (one object which is highlighted in the current marker colour) can be rotated through all objects of this type using the left and right arrow keys.</td>
</tr>
<tr>
<td>4.3.4</td>
<td>move</td>
<td>Move selected text objects.</td>
</tr>
<tr>
<td>4.3.5</td>
<td>copy</td>
<td>Copy selected text objects.</td>
</tr>
<tr>
<td>4.3.6</td>
<td>erase</td>
<td>Erase selected text objects.</td>
</tr>
<tr>
<td>4.3.7</td>
<td>modify</td>
<td>Modify selected text objects.</td>
</tr>
<tr>
<td>4.3.8</td>
<td>back</td>
<td>Transfer selected text objects to background</td>
</tr>
<tr>
<td>5.1</td>
<td>graphics</td>
<td>Indicator for current graphics colour</td>
</tr>
<tr>
<td>5.2</td>
<td>color</td>
<td>Indicator for current line thickness</td>
</tr>
<tr>
<td>5.3</td>
<td>thickness</td>
<td></td>
</tr>
<tr>
<td>5.4</td>
<td>TYPE</td>
<td>Graphics object types menu.</td>
</tr>
<tr>
<td>5.5</td>
<td>STYLE</td>
<td>Graphics style functions.</td>
</tr>
<tr>
<td>5.6</td>
<td>EDIT</td>
<td>Graphics edit functions</td>
</tr>
<tr>
<td>5.7</td>
<td>clear</td>
<td>Erase all graphics objects.</td>
</tr>
<tr>
<td>5.4.1</td>
<td>graphics</td>
<td>Add line</td>
</tr>
<tr>
<td>5.4.2</td>
<td>line</td>
<td>Add circle</td>
</tr>
<tr>
<td>5.4.3</td>
<td>circle</td>
<td>Add filled circle</td>
</tr>
<tr>
<td>5.4.4</td>
<td>disc</td>
<td>Add segment of a filled circle</td>
</tr>
<tr>
<td>5.4.5</td>
<td>segment</td>
<td>Add rectangle</td>
</tr>
<tr>
<td>5.4.6</td>
<td>rect</td>
<td>Add filled rectangle</td>
</tr>
<tr>
<td>5.4.7</td>
<td>bar</td>
<td></td>
</tr>
<tr>
<td>5.4.8</td>
<td>MORE</td>
<td>More object types</td>
</tr>
<tr>
<td>Menu tree command</td>
<td>Command / icon / indicator</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>5.4.9</td>
<td>MISC</td>
<td>Miscellaneous graphics functions</td>
</tr>
<tr>
<td>5.4.8.1</td>
<td><em>graphics</em></td>
<td>Add ellipse</td>
</tr>
<tr>
<td>5.4.8.2</td>
<td>ellipse</td>
<td>Add filled ellipse</td>
</tr>
<tr>
<td>5.4.8.3</td>
<td>fill. ell.</td>
<td>Add polygon</td>
</tr>
<tr>
<td>5.4.8.4</td>
<td>poly</td>
<td>Add circle arc</td>
</tr>
<tr>
<td>5.4.8.5</td>
<td>arc</td>
<td>Add filled polygon</td>
</tr>
<tr>
<td>5.4.8.6</td>
<td>fill. poly</td>
<td>Add line graph. Asks for the number of points and the corresponding values. This is a macro implying that all components can be edited as separate objects.</td>
</tr>
<tr>
<td>5.4.9.1</td>
<td><em>graphics</em></td>
<td>Add pie chart (macro).</td>
</tr>
<tr>
<td>5.4.9.2</td>
<td>graph</td>
<td>Add histogram (macro).</td>
</tr>
<tr>
<td>5.4.9.3</td>
<td>piechart</td>
<td>Rectangular grid. Supports accurate drawing. Not stored as an object. Disappears after Update or Clear</td>
</tr>
<tr>
<td>5.4.9.4</td>
<td>histo</td>
<td></td>
</tr>
<tr>
<td>5.4.9.5</td>
<td>grid</td>
<td></td>
</tr>
<tr>
<td>5.5.1</td>
<td><em>graphics</em></td>
<td>Indicator for current graphics colour.</td>
</tr>
<tr>
<td>5.5.2</td>
<td><em>color</em></td>
<td>Indicator for current line thickness.</td>
</tr>
<tr>
<td>5.5.3</td>
<td><em>thickness</em></td>
<td>Select graphics colour</td>
</tr>
<tr>
<td>5.5.4</td>
<td>color</td>
<td>Select line thickness</td>
</tr>
<tr>
<td>5.5.5</td>
<td>thickness</td>
<td>Select line style (solid, dashed, dotted)</td>
</tr>
<tr>
<td>5.5.6</td>
<td>linestyle</td>
<td>Select a fill pattern.</td>
</tr>
<tr>
<td>5.5.7</td>
<td>fill pattern</td>
<td></td>
</tr>
<tr>
<td>5.6.1</td>
<td><em>graphics</em></td>
<td>Indicator for current edit mode.</td>
</tr>
<tr>
<td>5.6.2</td>
<td>mode</td>
<td>Edit mode: object or field. In field mode the user can select a group of graphics objects inside a specifiable rectangle.</td>
</tr>
<tr>
<td>5.6.3</td>
<td>mode</td>
<td></td>
</tr>
<tr>
<td>5.6.4</td>
<td>move</td>
<td>Move selected graphics objects.</td>
</tr>
<tr>
<td>5.6.5</td>
<td>copy</td>
<td>Copy selected graphics objects.</td>
</tr>
<tr>
<td>5.6.6</td>
<td>erase</td>
<td>Erase selected graphics objects.</td>
</tr>
<tr>
<td>5.6.7</td>
<td>modify</td>
<td>Modify selected graphics objects.</td>
</tr>
<tr>
<td>5.6.8</td>
<td>back</td>
<td>Transfer selected text objects to background.</td>
</tr>
<tr>
<td>Menu tree command</td>
<td>Command / icon / indicator</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>5.6.9</td>
<td>zoom</td>
<td>Zoom selected graphics objects in or out.</td>
</tr>
<tr>
<td>6.1</td>
<td>photo</td>
<td>Grab a video picture</td>
</tr>
<tr>
<td>6.2</td>
<td>get</td>
<td>Picture edit functions</td>
</tr>
<tr>
<td>6.3</td>
<td>EDIT</td>
<td>Picture style functions</td>
</tr>
<tr>
<td>6.4</td>
<td>STYLE</td>
<td>Erase all picture objects.</td>
</tr>
<tr>
<td>6.5</td>
<td>clear</td>
<td>Set picture type: bi-level or 256 colours.</td>
</tr>
<tr>
<td>6.6</td>
<td>levels</td>
<td>Indicator for current picture type.</td>
</tr>
<tr>
<td>6.7</td>
<td>levels</td>
<td>Set threshold for bi-level grabbing</td>
</tr>
<tr>
<td>6.8</td>
<td>threshold</td>
<td>Indicator for current threshold</td>
</tr>
<tr>
<td>6.9</td>
<td>threshold</td>
<td>Clip picture(s). Define rectangular clip area and clip all picture objects which are (partly) within the area.</td>
</tr>
<tr>
<td>6.3.1</td>
<td>photo</td>
<td>De-clip picture(s). Inverse of clip: a clipped picture inside the specified area will be displayed in original size. Warning: once a clipped picture has been stored to disk, the de-clip function has no effect.</td>
</tr>
<tr>
<td>6.3.2</td>
<td>clip</td>
<td>Move selected pictures.</td>
</tr>
<tr>
<td>6.3.3</td>
<td>declip</td>
<td>Copy selected pictures.</td>
</tr>
<tr>
<td>6.3.4</td>
<td>move</td>
<td>Zoom picture(s). Pictures within the specified picture(s). Pictures within the specified field can be enlarged or reduced by a zoom-factor to be entered. Range: 12-800</td>
</tr>
<tr>
<td>6.3.5</td>
<td>copy</td>
<td>Erase selected pictures.</td>
</tr>
<tr>
<td>6.3.6</td>
<td>zoom</td>
<td>Modify selected pictures. Usable for bi-level pictures. Picture(s) within specified field get current fore- and background-colours.</td>
</tr>
<tr>
<td>6.4.1</td>
<td>photo</td>
<td>Adjust contrast</td>
</tr>
<tr>
<td>6.4.2</td>
<td>contrast</td>
<td>Indicator for current contrast level.</td>
</tr>
<tr>
<td>6.4.3</td>
<td>contrast</td>
<td>Adjust brightness.</td>
</tr>
<tr>
<td>6.4.4</td>
<td>brightness</td>
<td>Indicator for current brightness level.</td>
</tr>
<tr>
<td>6.4.5</td>
<td>brightness</td>
<td>Define foreground-colour (bi-level pictures).</td>
</tr>
<tr>
<td>Menu tree command</td>
<td>Command / icon / indicator</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>6.4.7</td>
<td>foregr. color</td>
<td>Indicator for current fore-ground colour</td>
</tr>
<tr>
<td>6.4.8</td>
<td>backgr. color</td>
<td>Define background-colour (for bi-level pictures).</td>
</tr>
<tr>
<td>6.4.9</td>
<td>backgr. color tablet</td>
<td>Indicator for current background-colour</td>
</tr>
<tr>
<td>7.1</td>
<td>colors</td>
<td>Indicator for current line drawing colour and fill colour.</td>
</tr>
<tr>
<td>7.2</td>
<td>thickness</td>
<td>Indicator for current line thickness.</td>
</tr>
<tr>
<td>7.3</td>
<td>enter</td>
<td>Input line drawing (using electronic writing tablet)</td>
</tr>
<tr>
<td>7.5</td>
<td>relocate</td>
<td>Relocate input. Define a point on screen; the first point of the next string gets the defined position; all subsequent input undergoes the same translation. Reset at leaving of the Enter mode.</td>
</tr>
<tr>
<td>7.6</td>
<td>color</td>
<td>Define colour</td>
</tr>
<tr>
<td>7.7</td>
<td>thickness</td>
<td>Define line thickness</td>
</tr>
<tr>
<td>7.8</td>
<td>EDIT</td>
<td>Line drawing edit functions</td>
</tr>
<tr>
<td>7.9</td>
<td>fill</td>
<td>Fill area with fill colour. Cursor defines point on screen as a centre for filling. When in fill mode the colour indicator is split up in the current writing and the current fill color, and the Color command controls the fill colour.</td>
</tr>
<tr>
<td>7.8.1</td>
<td>tablet</td>
<td>Clear line drawing objects.</td>
</tr>
<tr>
<td>7.8.2</td>
<td>clear</td>
<td>Copy selected line drawing objects.</td>
</tr>
<tr>
<td>7.8.3</td>
<td>copy</td>
<td>Move selected line drawing objects.</td>
</tr>
<tr>
<td>7.8.4</td>
<td>move</td>
<td>Erase selected line drawing objects.</td>
</tr>
<tr>
<td>7.8.5</td>
<td>erase</td>
<td>Modify selected line drawing objects.</td>
</tr>
<tr>
<td>7.8.6</td>
<td>modify</td>
<td>Transfer selected line drawing objects to background.</td>
</tr>
<tr>
<td>7.8.7</td>
<td>back</td>
<td>Indicator for currently selected macro library.</td>
</tr>
<tr>
<td>8.1</td>
<td>lib: name</td>
<td>Indicator for currently selected macro library.</td>
</tr>
<tr>
<td>Menu tree command</td>
<td>Command / icon / indicator</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>8.2</td>
<td>define</td>
<td>Define and store selected objects as a new macro. The program asks for a macro name. The user may also add an additional explanatory comment to support selection of stored macros.</td>
</tr>
<tr>
<td>8.3</td>
<td>call</td>
<td>Call specified macro and place the objects at a user definable position.</td>
</tr>
<tr>
<td>8.4</td>
<td>dir</td>
<td>Display directory of currently selected macro library. Includes the names and the explanatory comment lines.</td>
</tr>
<tr>
<td>8.5</td>
<td>change</td>
<td>Change macro library.</td>
</tr>
</tbody>
</table>

Remarks:

1) **In object mode** the object to be edited is specified by the arrow, PgDn and PgUp keys. In field mode the centre mouse button (or Ins key) toggles between 3 cursor types:
   a) Left hook / right hook - defines a rectangular field;
   b) Cross - defines a point in the area to be edited and correlates this with a second point to perform accurate moving or copying;
   c) Square - defines snap field in the edit area and correlates the end-point of a line in that field with the end-point of a line in a second snap field; in this way endpoints of lines of different objects (line, rectangle, polygon, bar) can be linked.

2) Selection of menu commands can be effected by the Enter key or the left mouse button. Input modes, menus, dialog boxes, etc., can be left by the Esc key or the right mouse button. At the main menu this results in a program quit menu (not included in the table).

3) Frequently used commands can be effected bypassing the menu structure by use of function keys (short cuts).

4) By function key F1 a help window is popped up to display information on the options of the current menu.

5) In communication mode the indicator 2.8 (and 2.9 in the case of the transmission of an object or page with a large data size) is always displayed independent of the current menu.
ACKNOWLEDGEMENTS

Soli Deo Gloria

The work presented in this thesis has largely been performed at the Electrical Engineering Department, University of the Netherlands Antilles, Curaçao. The National Aerospace Laboratory NLR, the Netherlands, enabled me to accomplish the necessary additional research on image compression and multimedia delay-reduction methods.

I am indebted to the University of the Netherlands Antilles for the opportunity and environment to commence and accomplish a major part of this work. Further, I would like to thank my former colleagues of the UNA Faculty of Engineering who supported and contributed to the work. Especially, J. Willems, R. Malacia, R. Lasten, and several students, for their enthusiastic support during the preparation and realization of the classroom experiment. A profound appreciation is felt for the inspiring discussions with Ir. H. Menckeberg, former UNA colleague and now director of Radio Holland International at Bonaire.

I would also like to thank the NLR for the opportunity to perform a number of important investigations as part of the NLR research program. Ir. A. van Dorp helped a lot with practical problems during the final part of the work. I am also grateful to my father in law Mr. G. P. Wijnbeek, B.A., for his valuable textual suggestions.

Special thanks are due to my thesis supervisor Prof. Dr. J. C. Arnbak for his invaluable suggestions and ever stimulating discussions throughout the entire work.

Finally, I would like to thank my wife, Gerdien, for her unflagging support, care and encouragement.
SAMENVATTING

Het onderwerp van dit proefschrift is de ontwikkeling van technieken ten behoeve van tele-educatie systemen, welke real-time afstandsonderwijs op professioneel niveau mogelijk maken. In vele landen komen situaties voor waarbij aanzienlijke delen van de bevolking verstoken zijn van enige vorm van hoger onderwijs door (bijvoorbeeld) geografische barrieres. Conventionele vormen van afstandsonderwijs zijn echter vaak minder effektief dan het reguliere contactonderwijs. Bovendien zijn de bestaande onderwijsinstituten meestal onvoldoende uitgerust om een dergelijke vorm van onderwijs te realiseren. Het is echter mogelijk om met behulp van real-time tele-educatie, waarbij gebruik gemaakt wordt van telecommunicatie en computertechnologie, een vorm van afstandsonderwijs te realiseren, waarmee contactonderwijs dicht benaderd wordt. Dit betekent dat deze methoden beter geïntegreerd kunnen worden in de organisatie en werkwijzen van bestaande instituten. Bovendien dragen de korte response-tijden en het sterk interactieve karakter bij tot een grotere effektiviteit.

Na een overzicht van bestaande tele-educatie systemen wordt het ontwerp beschreven van een systeem dat in het bijzonder geschikt is voor de extensie van educatieve faciliteiten van universiteiten naar geïsoleerde gemeenschappen. De studie is gericht op kleinschalige instituten zoals die veel voorkomen in derde-wereld landen. Een typisch voorbeeld hiervan is de Universiteit van de Nederlands Antillen (UNA) op Curaçao, dat als een casus is beschouwd. Na analyse van de eisen en randvoorwaarden is een PC-gebaseerd systeem ontwikkeld, waarbij zoveel mogelijk functies door middel van software zijn gerealiseerd. Hierbij zijn interactieve media toegepast zoals een vorm van video met een lage beeldfrequentie, audio, en de communicatie via een gemeenschappelijk scherm met behulp van computertekst en graphics, een elektronisch schrijftablet, en een document-scanner.

Er wordt een snelle en effektive coderingsmethode beschreven voor de overdracht van sekwenties van video-beelden. Deze techniek, die geschikt is voor toepassing in microcomputersystemen zonder gebruikmaking van specifieke hardware, leidt tot relatief korte overdrachtstijden, ook in situaties waarbij de bandbreedte van het communicatiekanaal een beperking vormt. Een onderdeel van deze methode is een nieuwe benadering voor entropiecodering. Modified Tunstall Coding (MTC) is een verbetering van het door Tunstall voorgestelde compressie-algoritme voor de omzetting van reeksen bronsymbolen met een variabele lengte naar codewoorden met een vaste lengte. MTC is
een aantrekkelijke vorm van entropiecodering bij bronnen met een beperkt alfabet en een lage entropie, in het bijzonder bij implementatie in software.

Vervolgens wordt een aantal verbeteringen voorgesteld voor het digitaal coderen en versturen van handgeschreven tekst bij toepassing ten behoeve van tele-educatie. Het betreft een efficiente coderingstechniek, een additionele mogelijkheid om oppervlakken van getekende objekten in te kleuren, en een verbeterde methode voor de bewerking van bestaande gegevens. Real-time tele-educatie vereist dat de overdrachtstijden van visueel materiaal (bijvoorbeeld multimedia documenten) minimaal zijn. Daarvoor is een methode ontwikkeld ("pretransfer") waarbij de fluktuerende overcapaciteit van het communicatiekanaal tijdens de sessie wordt gebruikt voor het automatisch versturen van deze objekten, vóór het gewenste moment van weergave.

De genoemde technieken en oplossingen zijn geïntegreerd in het ontwerp van een realtime tele-educatie systeem, dat geschikt is voor smalle bandbreedtes. De ontwikkeling en de implementatie van dit interactieve multimediaysysteem worden beschreven. Een initieel prototype is gebruikt voor de uitvoering van een experiment in een werkelijke onderwijs situatie op de UNA. Uit de evaluatie-resultaten kan geconcludeerd worden dat het ontwikkelde multimedia-concept aan de gestelde eisen voldoet. Tele-educatie in deze vorm blijkt veelbelovend te zijn om effektief afstandsonderwijs te realiseren in de omschreven doelsituaties.