Fusion of Perceptions in Architectural Design

Ozer Ciftcioglu[1], *Michael S. Bittermann*[2]
Delft University of Technology, The Netherlands
http://bk.tudelft.nl/en/research/research-projects/computational-intelligent-design
[1]*o.ciftcioglu@tudelft.nl*, [2]*m.s.bittermann@tudelft.nl*

Abstract. *A method for fusion of perceptions is presented. It is based on probabilistic treatment of perception, where perception quantifies the chance an unbiased observer sees an environmental object, and the associated probability can be interpreted as degree of awareness for the object. The approach uniquely accounts for the fact that final realization or remembrance of a scene in the brain may be absent or elusive, so that it is subject to probabilistic considerations. For objects that are to be perceived from multiple viewpoints, such as a sculpture in a museum, or a building in its urban context, the probabilistic approach uniquely defines the fusion of perceptions. This is accomplished by carrying out the probabilistic union of events. The computation is presented together with its geometric implications, which become rather intricate for multiple observers, whereas the computation is straight forward. The method is exemplified for two applications in architectural design at different scales, namely interior and urban design, indicating the generic nature as well as the large application potential of the method.*
Keywords. *Perception; vision modeling; architectural design; evolutionary search.*

INTRODUCTION

Perception, and in particular visual perception, is an interdisciplinary concept taking an important place in many diverse applications. These range from design of objects and spaces, for which perceptual qualities are aimed (Bittermann and Ciftcioglu, 2008), to robotics where a robot moves based on perception (Ciftcioglu et al., 2006a; Bülthoff et al., 2007). However, although visual perception has been subject to scientific study for over a century, e.g. see Wertheim (1894), it is interesting to note that it remained mysterious what perception precisely is about, while it eluded mathematical modeling until very recently. Many approaches to perception, in particular in the domain of psychology and neuroscience, are based on experiment, while underlying theoretical models or hypotheses are either simplistic, ambiguous or even absent (Treisman and Gelade, 1980; O'Regan et al., 2000; Treisman, 2006), so that gaining insight into the nature of human perception from the experiments remains minimal. However, considering that the perception phenomenon is due to brain processing of retinal photon-reception, it should be clearly noted that the phenomenon is highly complex. That is, the same experimenter may have different perceptions of the same environment at different times, depending on the complexity of the environment, psychological state, personal preferences and so on, not to mention different vantage points. Due to the complexity of the brain processes and diversity of environments subject to visual perception, the empiric approaches to perception yielded merely rudimentary understanding of what perception is. Although some verbal definitions of the concept are presented in the literature, e.g. (Gibson, 1986; Palmer, 1999; Foster, 2000; Smith, 2001) due to excessive ambiguity of the linguistic expressions, these are not to be converted to precise or even more unambiguous mathematical expressions.

Computational approaches addressing some perception aspects have been proposed by Marr (Marr, 1982) whose prescription is to build computational theories for perceptual problems before modeling the processes which implement the theories. Explicitly, different visual cues are computed in separate modules and thereafter only weakly interact with each other, where each module separately estimates scene properties, such as depth and surface orientation, and then the results are combined in some way. These works can be termed as image processing based approaches, and they are deterministic in nature, starting from simulation of retinal data acquisition. The retinal photon-reception certainly is the first stage in the time sequence of the processing in the visual system, and it might be dealt with by means of an image specified as a two-dimensional matrix. However, the ensuing neural processes are highly complex, so that retinal image does not imply that all the information in the scene is registered in the human brain and remembered shortly afterwards. Only part of the visual information is remembered. For instance, it is a common experience that when we look at a scene, we are not aware of the existence of all objects the scene comprises. This is easily verified for scenes where the number of objects exceeds about seven objects.

In this work a probabilistic approach is adopted for perception, where perception is considered a whole process from the stimulus coming from the scene to mental realization in the brain. In other words, all complex processes, e.g. image formation on the retina, processes in the visual cortex in the brain, and final realization of 'seeing' is modeled as a single probabilistic event, where 'seeing' in that probabilistic description is considered to be perception, where remembrance is a matter of probability. The final realization or remembrance of the scene in the brain may be absent or elusive, which is subject to probabilistic considerations. This approach has been described and its validity demonstrated (Ciftcioglu et al., 2006b; Bittermann and Ciftcioglu, 2008).

This probabilistic approach is unique in the sense that the perception refers to human perception. In the field of computer vision perception is considered to be a mere image processing and ensuing pattern recognition process, where Bayesian methods are appropriate (Knill et al., 2008; Knill and Richards, 2008; Yuille and Bulthoff, 2008). Bayesian approach is to characterize the information about the world contained in an image as a probability distribution which characterizes the relative likelihoods of a viewed scene being in different states, given the available image data. The conditional probability distribution is determined in part by the image formation process, including the nature of the noise added in the image coding process, and in part by the statistical structure of the world. The Bayes's rule provides the mechanism for combining these two factors into a final calculation of the posterior distribution. This approach is based on Bayes formula

$$p(s \mid i) = \frac{p(i \mid s)\, p(s)}{p(i)} \tag{1}$$

Here *s* represents the visual scene, the shape and location of the viewed objects, and *i* represents the retinal image. *p(i|s)* is the likelihood function for the scene and it specifies the probability of obtaining image *i* from a given scene *s*. *p(s)* is the prior distribution which specifies the relative probability of different scenes occurring in the world, and formally expresses the prior assumptions about the scene structure including the geometry, the lighting and the material properties. *p(i)* can be derived from *p(i|s)* and *p(s)* by elementary probability theory. Namely

$$p(i) = p(i \mid s)\, p(s) + p(i \mid \bar{s})\, p(\bar{s}) \tag{2}$$

so that (1) becomes

$$p(s \mid i) = \frac{p(i \mid s)\, p(s)}{p(i \mid s)\, p(s) + p(i \mid \bar{s})\, p(\bar{s})} \tag{3}$$

The posterior distribution *p(s|i)* is a function giving the probability of the scene being *s* if the observed image is *i*. Bayesian approach is appropriate for

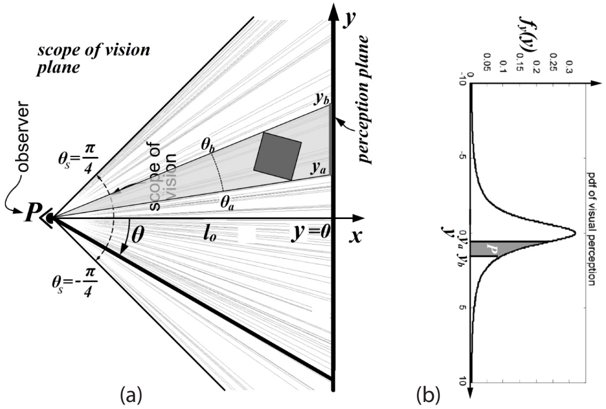


Figure 1
Plan view of the basic geometric situation of perception; P represents an observer's point, viewing an object (a); probability density function characterizing perception along y direction for l_o=2 (b).

computer vision, because for human *p(i|s)* is almost clearly known, that is *p(i|s)*=1. Consequently, *p(i|$\bar{s}$)*=0 and from equation (3)

$$p(s \mid i) = \frac{1 \times p(s)}{1 \times p(s) + 0 \times p(\bar{s})} = 1 \tag{4}$$

which is independent of the probabilistic uncertainties about the scene. This means, as the *p(i|s)* is definitive for human recognizing a scene, *p(s|i)* is also definitive, being independent of *p(s)* which is the prior assumptions about the scene structure including the geometry, the lighting and the material properties. The effectiveness of Bayes for machine vision is due to its recursive form, providing improved estimation as the incoming information is sustained.

The organization of the paper is as follows. In the *modeling human perception* section a vision model is established. In the *perception from multiple viewing positions* section, the fusion of perceptions from multiple viewpoints is derived. In the section *experiments,* two experiments demonstrating the fusion of perceptions in architectural design are presented, and this section is followed by *conclusions*.

MODELLING HUMAN PERCEPTION

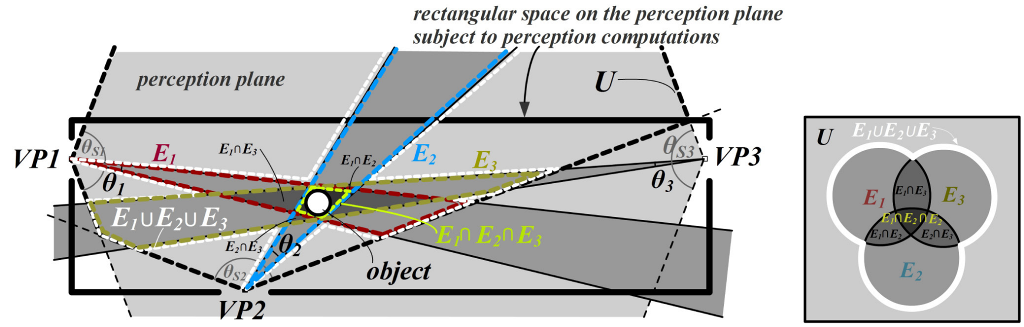
In the human perception an object is visually seen, but its remembrance is subject to some degree via probabilistic considerations. This is described elsewhere (Ciftcioglu et al., 2006b; Bittermann and Ciftcioglu, 2008) and briefly mentioned as follows. We consider a basic geometric situation as shown in Figure 1a. For a visual scope $-\pi/4 \leq \theta \leq \pi/4$ the probability density characterizing perception along the *y*-direction is shown in Figure 1b for $l_o=2$ and given by

$$f_y(y) = \frac{2}{\pi} \frac{l_o}{(l_o^2 + y^2)} \quad (-l_o \leq y \leq +l_o) \tag{5}$$

The probability density with respect to *q* is given by $f_\theta(\theta)=1/\theta_S$, where $\theta_S=\pi/2$. The one-dimensional perception of an object spanning from arbitrary object boundaries *a* and *b* on the *y*-axis is obtained by

$$P_y = \int_{ya}^{yb} f_y(y)\, dy \tag{6}$$

yielding perception as an event being subject to probabilistic computation. For the case of perception of an object by a single human observer the computation is accomplished always by (6) when the projection of the object is considered as one-dimensional along a line. The same computation can be valid for three-dimensional objects, provided we consider the projection of the object on a plane. In this case, the same formulation can be used twice for each respective orthogonal dimension of the plane in the form of product of the two probability densities integrated over the projected area on the plane.



(a) (b)

Figure 2
Perception events E_1, E_2, and E_3 respectively denoting perception of an object from three viewpoints VP_1, VP_2, VP_3; the union of the events is indicated by the white dashed line (a); Venn diagram corresponding to the perception events in Figure 2a (b).

PERCEPTION FROM MULTIPLE VIEWING POSITIONS

In many occasions an object is subject to perception from multiple viewing positions, either by the same observer or by multiple observers. That is, perceptions from different viewing positions are subject to fusion. As the perception is expressed in probabilistic terms, the union of different perception events is subject to probabilistic computation. Requirements with respect to perception from multiple viewing positions can occur in many practical applications. To demonstrate fusion of perceptions we restrict the study to two basic examples. It is noted that they may not be important depending on the particular design problem; however the examples are simple in order to clearly explain the method. The same method can be applied in more complex tasks, such as courtroom design (Bhatt et al., 2011), auditorium design, office design, as well as urban design. In the first case study we consider an exhibition gallery environment, where there are several entrances to a gallery space, and we are wondering what the best position to place an object is, so that perception of the object is maximized. In the second application we are considering an urban environment, where a building will be erected that will be seen from a number of prominent viewing positions. We are interested to obtain the perception of the different parts of the future building as fusion of perceptions from these viewing positions. In the latter case study this is to identify which part of the building is most conspicuous, in order to determine for instance the building entrance that should preferably be positioned. Consequently it will be easily noticed.

A scene subject to investigation as exemplary case is shown in Figure 2a, with the three perception events E_1, E_2, and E_3. The figure shows a plan view of the space and the location of an object subject to perception assessment and optimal positioning. The object is subject to perception from the three viewing positions VP_1, VP_2, and VP_3, where it respectively subtends the angle domains θ_1, θ_2, and θ_3 as seen in the figure. The dashed lines in the figure indicate the boundaries of the observer's visual scope at the respective viewing positions spanning the angles θ_{S1}, θ_{S2}, and θ_{S3}. Figure 2b shows a Venn diagram corresponding to the perception situation in Figure 2a. In the case of perceiving an object from several viewing positions this corresponds to the probabilistic union of the perceptions, which is obtained by $P(E_1 \cup E_2 \cup E_3)=P(E_1)+P(E_2)+P(E_3)-P(E_1 \cap E_2)-P(E_1 \cap E_3)-P(E_2 \cap E_3)+P(E_1 \cap E_2 \cap E_3)$, as this is seen from Figure 2b. It is noted that the events $P(E_1)$, $P(E_2)$, and $P(E_3)$ are independent. In the three dimensional perception case θ_1, θ_2, and θ_3 become solid angles Ω_1, Ω_2, and Ω_3 and the scopes θ_{S1}, θ_{S2}, and θ_{S3} become solid angles Ω_{S1}, Ω_{S2}, and Ω_{S3}.

EXPERIMENTS

Computer experiments are carried out, where $P(E_1)$, $P(E_2)$, and $P(E_3)$ are obtained by probabilistic ray tracing, so that a three-dimensional object is subject to perception measurement without need for projection to a plane as shown in Figure 1a. That is, the solid perception angle Ω subtended by the object, as well as the solid angle Ω_S, which defines the observer's visual scope, are simulated by vision rays that are sent in random directions within the three dimensional visual scope. The randomness in terms of unit Ω is characterized by $f_\Omega(\Omega)=1/\Omega_S$, conforming to the uniform pdf $f_\theta(\theta)=1/\theta_S$ that models the unbiased observer in the case of perception of an object that is contained in the scope of vision plane, as seen in Figure 1a. In the experiments the number of vision rays is denoted by n_v. An object within the visual scope will be hit by a number of vision rays n_p, and these rays are termed perception rays. The perception of the object is given by $P=n_p/n_v$.

Experiment Nr. 1

The first experiment concerns a basic issue in an architectural design, namely positioning an object, so that its perception from several viewing positions is maximized in the sense that the object will be perceived well at least from one of the relevant viewpoints. This issue is exemplified by means of positioning a sculpture in a museum space having several entrances; namely the space has three doors, where the relevant viewing positions are located denoted by VP_1, VP_2, and VP_3. The problem is to position the sculpture in the space, so that the visitors entering the space from either door will notice the object. The problem is to maximize the union of the perceptions from the three viewpoints $P(E_1 \cup E_2 \cup E_3)$, while at the same time the sculpture positioned at point x should not obstruct entrance to the room from either door. The latter constraint is formulated by the condition $\|x-x_o\| \geq 3$, where x_o is the position of each viewing position. The maximization is carried out by the method of random search, accomplished through the method of genetic algorithm. Genetic algorithm is a stochastic optimization method from the domain of computational intelligence. The algorithm starts from a number of random solutions referred to as members of a population. Each member satisfies the objective function to some degree, which is termed fitness. In the algorithm population members with a comparatively high fitness will be favored over solutions with low fitness, by giving the former a higher chance to remain in the population and to produce new solutions by combining fit solutions. The combination among solutions is referred to as crossover operation, and it is carried out among pairs of population members referred to as parents. Crossover entails that the parameters constituting a parent are treated as binary strings, and portions of the strings are exchanged among the two solutions to create new solutions with features from both parents. This process is repeated for several iterations, and due to the probabilistic favoring of fit solutions, eventually optimal solutions appear in the population (Goldberg, 1989; Zalzala and Fleming, 1997).

The resulting best solution after 40 generations is shown in Figure 3a in a plan view and in Figure 3b in perspective view, where the perception rays are seen. The circles in the figures mark the boundaries at 3.0 m distance from the doors. In Figure 3c-e the space is shown from the respective viewing position VP_1, VP_2, and VP_3. The best position of the sculpture is at the edge of the circle in front of viewing position VP_1. This position has the highest union of perceptions in the feasible region, namely P_U=.307. This is composed of the perceptions P_1=.157 at VP_1, P_2=.063 at VP_2, and P_3=.048 at VP_3.

For comparison the second best position is shown in Figure 4, namely the perceptive plan view in Figure 4a, perspective perceptive view in Figure 4b, and the perceptive views from VP_1, VP_2, and VP_3 in Figure 4c-e respectively. The union of the perceptions P_U=.255, that is 17% lower compared to the best solution in Figure 3. The union is composed of the perceptions P_1=.072 at VP_1, P_2=.152 at VP_2, and P_3=.033 at VP_3. The results demonstrate a common design knowledge, namely when one aims to maximize the perception of an object in a space with several possible viewing positions, it is preferable to position the object to have a high perception for at least one of the possible positions, for that matter VP_1, rather than having several moderate percep-

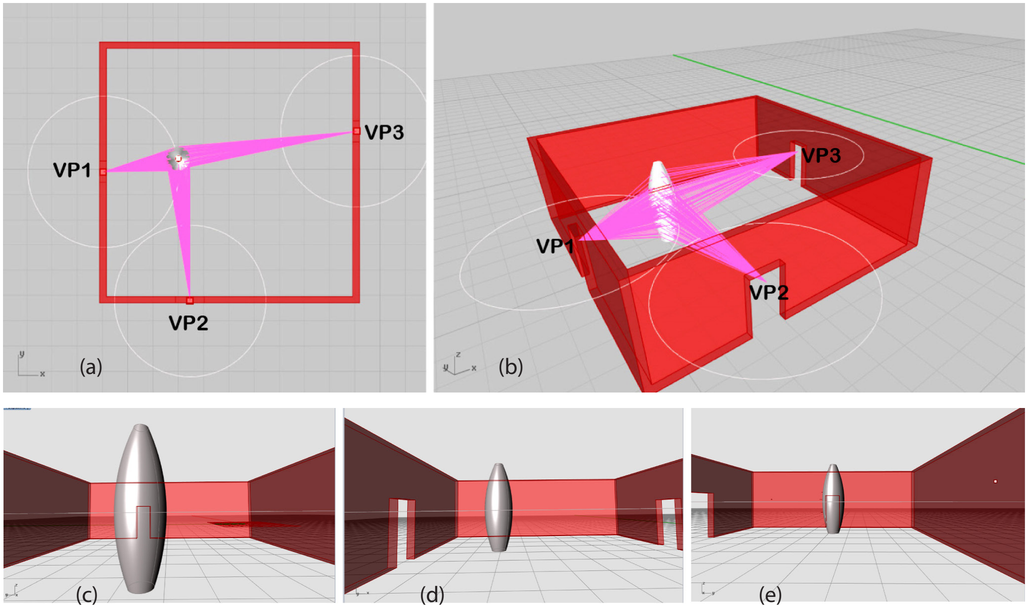


Figure 3
Best position for the sculpture in a plan view (a); in a perceptive view (b); for VP_1, P_1=.157 (c); for VP_2, P_2=.063 (d); for VP_3, P_3=.048 (e).

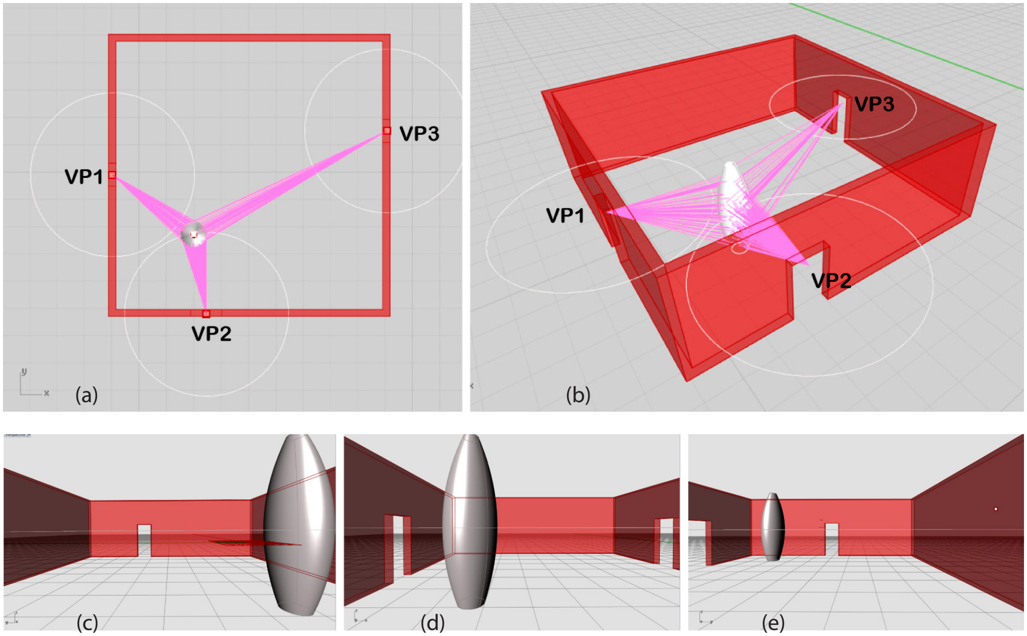


Figure 4
Second best position for the sculpture in a plan view (a); in a perceptive view (b); for VP_1, P_1=.072 (c); forVP_2, P_2=.152 (d); for VP_3, P_3=.033 (e).

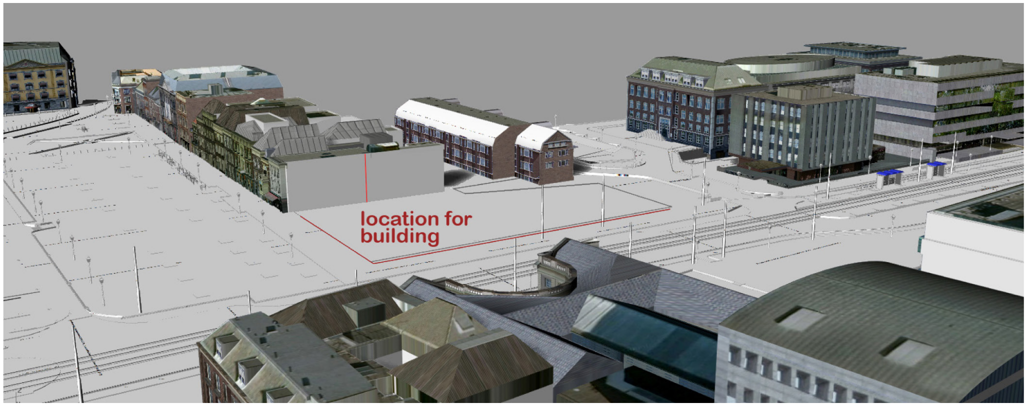


Figure 5
Location in an urban scene, where a new building is subject to perception considerations.

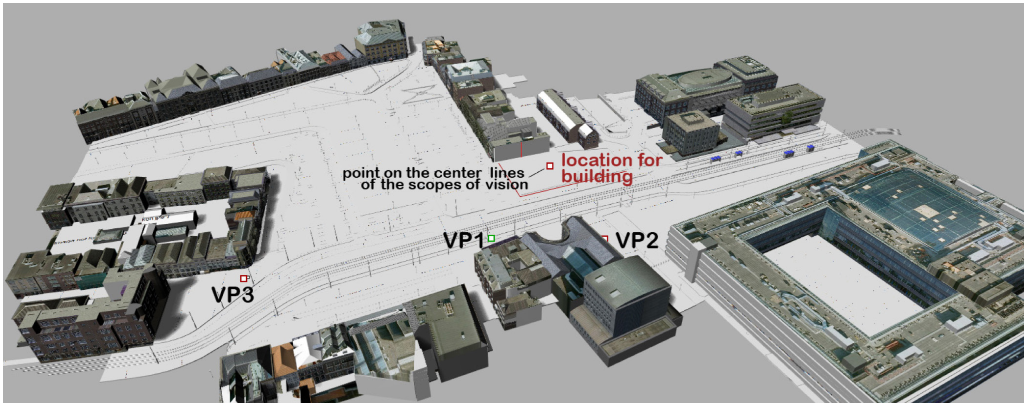


Figure 6
Zoomed out rendering of the urban scene in Figure 5, where a new building is subject to perception considerations from three viewpoints.

tions, i.e. without any outstandingly high one. The lower perceptions in Figure 4c demonstrate the implications of the Cauchy function in Figure 1b, where deviation from the frontal direction for an object, in particular at a near distance from the observer, yield reduction in probability density, i.e. visual attention is diminished in this case.

Experiment Nr. 2

A second experiment concerns the perception of a building in an urban context from three viewpoints that are prominent locations in the surrounding of the building. The location in an urban scene, where a new building is subject to perception considerations, is seen in Figure 5. The zoomed out rendering of the scene in Figure 5 is shown in Figure 6, where the three viewing positions VP_1, VP_2, and VP_3 are indicated. Figure 7 schematically shows the floor plan of the urban situation, as well as the perception cones and vision scopes belonging to the viewing positions, which are the endpoints of streets entering to a square where the building is located. Figure 7b shows random vision rays having uniform pdf with respect to the vision angle modeling visual scopes for three viewing positions. Figure 7c shows those rays among the vision rays that hit the building subject to perception, for perception computation. The results from the perception fusion for the respective building envelope portions are shown Figure 8, where the numbers display the fused perception associated with the respective portion. From the analysis it is seen that the part of the envelope that is most intensely perceived from the three viewpoints, is the area in front of VP_2, while the second most intense part is the part of the building corner oriented towards VP_1, which is expected considering the influence of the distance l_o in the perception computations in (5). The information obtained from perception fusion is of relevance for a designer determining formal and functional details of the envelope, for instance determining the position of entrance during conceptual design. Figure 9 shows the fused perceptions of the building envelope from the three viewpoints per envelope element with a vision scope that is 20% narrower compared to Figure 8.

CONCLUSIONS

A method for fusion of perceptions is presented and demonstrated with two examples from architectural design. The probabilistic treatment, where perception quantifies the chance that an unbiased observer notices an environmental object, is accomplished through fusion of perceptions. The method of quantified union of perceptions has been an unresolved issue up till now, that is resolved in this presentation. The fusion by probabilistic union yields significant information for designers. With the presented approach an object is to be perceived from several viewpoints at the same time. Such abstraction is necessary, since the precise analysis of the perceptions is a formidable issue due to abundant visual scene information. The use of perception fusion as constrained design objective has been demonstrated by coupling the method with a probabilistic evolutionary algorithm performing the constraint optimization. The combination of the two probabilistic methods is a powerful tool for designers as it permits treatment of architectural design to be highly constrained and involving many perception related demands. Although the examples presented are rather basic, the method is generic and yields highly appreciable scoring executions in diverse ap-

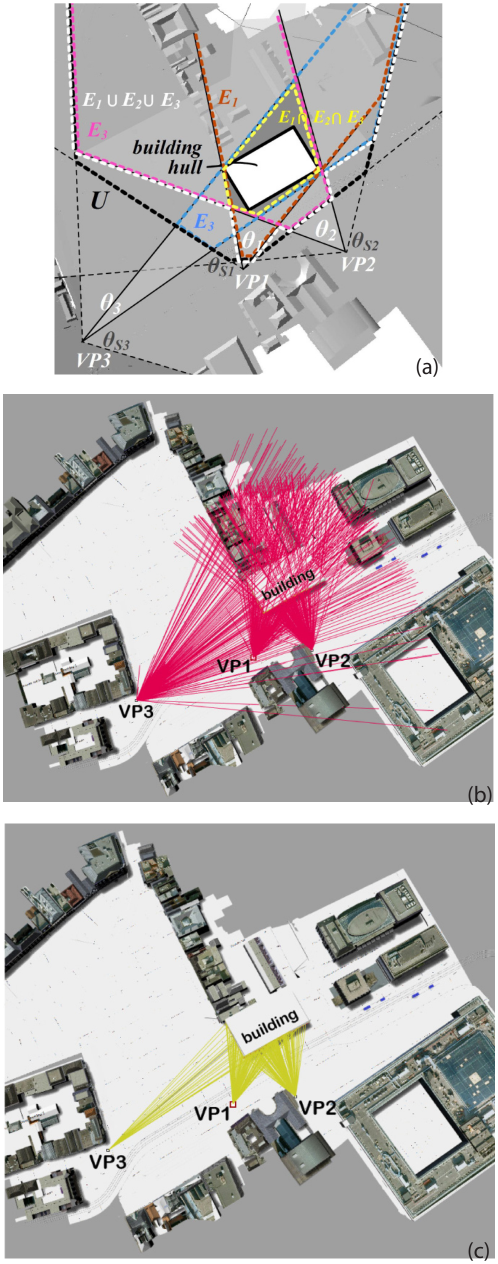


Figure 7
Scheme of an urban situation, where a building is subject to perception analysis from three viewpoints in plan view (a); random vision rays with uniform pdf w.r.t. the vision angle modeling visual scopes for three viewing positions VP_1, VP_2, and VP_3 (b); the rays among the vision rays that hit the building subject to perception (c).

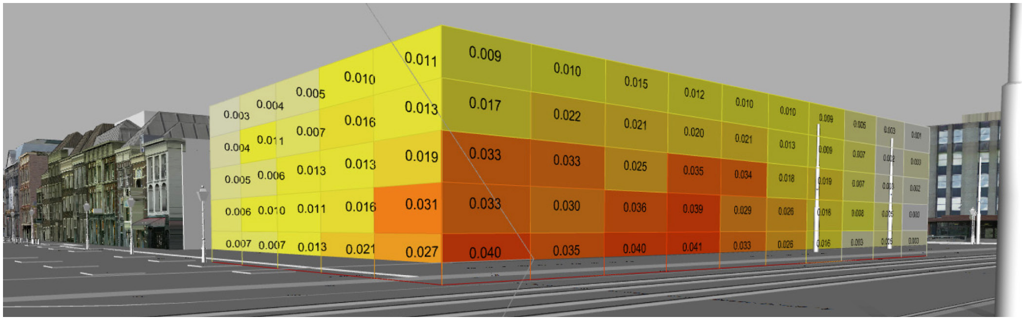


Figure 8
Fused perceptions of the building envelope from the three viewpoints per envelope element.

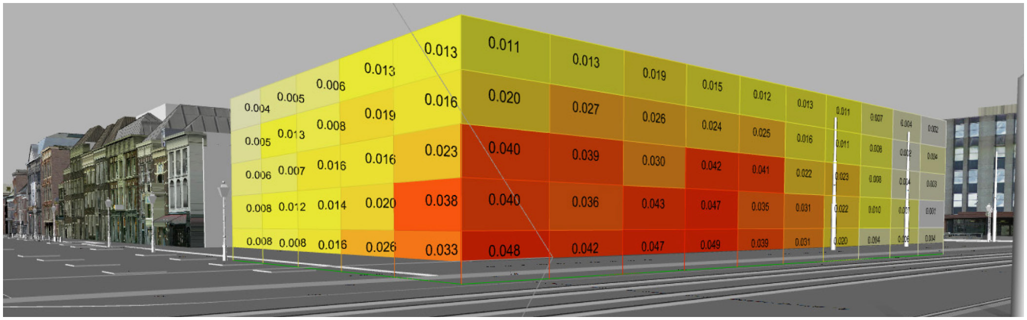


Figure 9
Fused perceptions of the building envelope from the three viewpoints per envelope element with a vision scope that is 20% narrower compared to Figure 8.

plications in the areas where perception plays a role, such as architecture, urbanism, interior and industrial design, as well as robotics.

ACKNOWLEDGEMENT

Technical design assistance by Architect Paul de Ruiter, providing us with the scene presented in the second experiment is gratefully acknowledged.

REFERENCES

Bhatt, M, Hois, J and Kutz, O 2011, 'Ontological Modelling of Form and Function for Architectural Design', *Applied Ontology*, pp. 1-32.

Bittermann, MS and Ciftcioglu, Ö 2008, 'Visual perception model for architectural design', *Journal of Design Research*, 7(1), pp. 35-60.

Bülthoff, H, Wallraven, C and Giese, M 2007, 'Perceptual robotics', in B Siciliano and O Khatib (eds), *The Springer Handbook of Robotics*, Springer, pp. 1481-1495.

Ciftcioglu, Ö, Bittermann, MS and Sariyildiz, IS 2006a, 'Studies on visual perception for perceptual robotics', *ICINCO 2006 - 3rd Int. Conf. on Informatics in Control, Automation and Robotics*, Setubal, Portugal, pp. 468-477.

Ciftcioglu, Ö, Bittermann, MS and Sariyildiz, IS 2006b, 'Towards computer-based perception by modeling visual perception: a probabilistic theory', *2006 IEEE Int. Conf. on Systems, Man, and Cybernetics, Taipei*, Taiwan, pp. 5152-5159.

Foster, J 2000, *The Nature of Perception*, Oxford University, Oxford.

Gibson, JJ 1986, *The Ecological Approach to Visual Perception*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Goldberg, DE 1989, *Genetic Algorithms*, Addison Wesley, Reading, MA.

Knill, DC, Kersten, D and Mamassian, P 2008, 'Implications of a Bayesian formulation for visual information for processing for psychophysics', *Perception as Bayesian Inference*, Cambridge, Cambridge, pp. 239-286.

Knill, DC and Richards, W 2008, *Perception as Bayesian Infer-*

ence, Cambridge University, Cambridge, UK.
Marr, D 1982, Vision, Freeman, San Francisco.
O'Regan, JK, Deubel, H, Clark, JJ and Rensink, RA 2000, 'Picture changes during blinks: looking without seeing and seeing without looking', *Visual Cognition*, 7(1-3), pp. 191-211.
Palmer, SE 1999, *Vision Science*, MIT, Cambridge, MA.
Smith, D 2001, *The Problem of Perception*, Harvard University, Cambridge, MA.
Treisman, AM 2006, 'How the deployment of attention determines what we see', *Visual Cognition*, 14(4), pp. 411-443.
Treisman, AM and Gelade, G 1980, 'A feature-integration theory of attention', *Cognitive Psychology*, 12(1), pp. 97-136.
Wertheim, T 1894, 'Ueber die indirekte Sehschaerfe', *Z Psychol Physiol Sinnesorg*, 7, pp. 172-189.
Yuille, AL and Bulthoff, HH 2008, 'Bayesian decision theory and psychophysics' in DC Knill and W Richards (eds), *Perception as Bayesian Inference*, Cambridge University, Cambridge, UK, pp. 123-161.
Zalzala, AMS and Fleming, PJ 1997, *Genetic Algorithms in Engineering Systems*, IEE Control Eng., Series 55, Cambridge University, New York.