Towards Computer-Based Perception by Modeling Visual Perception: A Probabilistic Theory

O. Ciftcioglu, Senior Member, IEEE, M. Bittermann, and Sevil Sariyildiz

Abstract—Studies on computer-based perception by vision modelling are described. The visual perception is mathematically modelled where the model receives and interprets visual data from the environment. The perception is defined in probabilistic terms so that it is in the same way quantified. Human visual perception mimicked by means of a computer is an important step in cybernetics as this is, generally speaking, one of the goals of cybernetics. At the same time, the measurement of visual perception is made possible in real-time. From the visual perception, some other derivates of it can be computed. One example is the visual openness which can be used for the movement of an autonomous robot. As to another application, mention may be made to spatial design, in building and construction engineering. The paper describes the novel probabilistic theory of visual perception and investigates various properties of it, via the vision model established. The computer experiments are carried out by means of virtual agent in virtual environment demonstrating the verification of the theoretical considerations being presented. At the same time, experimental studies are presented as to the derivates of visual perception demonstrating the far reaching implications of the studies.

I. INTRODUCTION

Visual perception is one of the important subjects of cybernetics. As human gets about eighty per cent of environmental information by visual perception, it is easy to understand the importance of it, if one endeavors to integrate the environmental information into a machine-based system. To achieve this goal, it is clear to realize that the visual perception should be quantified to feed it to computer, rather than remaining in the abstract domain and merely to comprehend the concept and dealing with some verbal statements. In the case of human, the interaction with the environment is done via the light photons emitted from a light source, scattered from the object and eventually reaching the retina. In the computer based vision process, the virtual agent or a robot, emits rays and the interaction of these rays with environment is registered as they backscattered to agent or robot again. Today such systems are quite well developed as they are referred to as 3D scanners. For the sake of clarity in the experimental part of the work, a virtual agent is used as a representative of a human, who moves through a spatial environment experiencing continuous visual perception. Next to this, also it makes assessment about the visual openness derivative of the visual perception so that it can make autonomous movement and in the meanwhile make perception related measurements. The organisation of the paper is as follows. Chapter two describes the theoretical considerations about the visual perception. Chapter three gives the details of the computer experiments for the verification of the theory. This is followed by conclusions.

II. A PROBABILISTIC THEORY FOR VISUAL PERCEPTION

A. The basic Visual Perception Model

We start with the basics of the perception process with a simple yet a fundamental visual geometry. This is shown in figure 1.

![Fig. 1. The geometry of visual perception from a top view where P represents the position of eye, looking at a vertical plane with a distance \( l_o \) to the eye; \( f(z) \) is the probability density function](image)

In figure 2, the observer is facing and looking at a vertical plane from the point denoted by P. By means of looking action the observer pays visual attention equally to all locations on the plane in the first instance. That is, the observer visually experiences all locations on the plane without any preference for one region over another. Each point on the plane has its own distance within the observer’s scope of sight which is represented as a cone. The cone has a solid angle denoted by \( \theta \). The distance of a point on the plane and the observer is denoted by \( x \) and the distance between the observer and the plane is denoted by \( l_o \). Since the elements of visual openness perception are determined...
via the associated distance, it is straightforward to proceed to express the distance of visual perception in terms of θ. From figure 2, this is given by

\[ x = \frac{l_o}{\cos(\theta)} \]  

(1)

Since we surmise the observer pays visual attention equally to all locations on the plane in the first instance, the probability of getting attention for each point on the plane is the same so that the associated probability density function (pdf) is uniformly distributed. This posits ensures that there is no visual bias at the beginning of visual perception as to the differential visual resolution angle \( d\theta \). Assuming the scope of sight is defined by the angle \( \theta = \pm \pi/4 \), the pdf \( f_\theta \) is given by

\[ f_\theta = \frac{1}{\pi/2} \]  

(2)

Since \( \theta \) is a random variable, the distance \( x \) in (1) is also a random variable. The pdf \( f_x(x) \) of this random variable is computed as follows.

To find the pdf of the variable \( x \) denoted \( f_x(x) \) for a given \( x \) we consider the theorem on the function of random variable and solve the equation [1] Papoulis

\[ x = g(\theta) \]  

(3)

for \( \theta \) in terms of \( x \). If \( \theta_1, \theta_2, \ldots, \theta_n, \ldots \) are all its real roots,

\[ x = g(\theta_1) = g(\theta_2) = \ldots = g(\theta_n) = \ldots \]

Then

\[ f_x(x) = \frac{f_\theta(\theta_1)}{|g'(\theta_1)|} + \frac{f_\theta(\theta_2)}{|g'(\theta_2)|} + \ldots + \frac{f_\theta(\theta_n)}{|g'(\theta_n)|} + \ldots \]  

(4)

Clearly, the numbers \( \theta_1, \theta_2, \ldots, \theta_n, \ldots \) depend on \( x \). If, for a certain \( x \), the equation \( x = g(\theta) \) has no real roots, then \( f_x(x) = 0 \).

According the theorem above,

\[ g'(\theta) = \frac{l_o \sin(\theta)}{\cos^2(\theta)} \]  

(5)

Between \( \theta = -\pi/4 \) and \( \theta = +\pi/4 \),

\[ g(\theta) = \frac{l_o}{\cos(\theta)} \]  

(6)

has two roots, which are equal and given by

\[ \theta_{l_o} = \arccos\left(\frac{l_o}{x}\right) \]  

(7)

Using (7) in (5), we obtain

\[ g'(\theta) = \frac{x\sqrt{x^2 - l_o^2}}{l_o} \]  

(8)

Substituting (2), (7) and (8) into (4), we obtain

\[ f_x(x) = \frac{4}{\pi} \frac{l_o}{x \sqrt{x^2 - l_o^2}} \]  

(9)

for the interval \( l_o \leq x \leq \sqrt{2} l_o \). For this interval, the integration below becomes

\[ \int_{l_o}^{\sqrt{2} l_o} f_x(x) \, dx = \frac{4}{\pi} \int_{l_o}^{\sqrt{2} l_o} \frac{l_o}{x \sqrt{x^2 - l_o^2}} \, dx = 1 \]  

(10)

as it should be as pdf. The sketch of \( f_x(x) \) vs \( x \) is given in figure 2.

\[ f_x(x) = \frac{4}{\pi} \frac{l_o}{x \sqrt{x^2 - l_o^2}} \]  

\[ 0 \leq x \leq l_o \]

As to (9), two observations are due. Firstly, it is interesting to note that for the plane geometry in figure 1, the visual perception is sharply concentrated close to \( \theta \approx 0 \), that is perpendicular direction to the plane. This striking result is in conformity with the common human experience as to visual perception. Namely, for this geometry the visual perception is strongest along the axis of the cone of sight relative to the side directions. This is simply due to the fact that, for the same differential visual resolution angle \( d\theta \), one can perceive visually more details on the infinite plane in the perpendicular direction as this is sketched in figure 3.
Fig. 3. Sketch explaining the relative importance of the viewing direction for visual perception.

Secondly, the visual perception is given via a probability density at a point. If we consider the stimulus of perception is due to the light photons, it is the relative number of photons as stimulus at infinitesimally small interval, per unit length. Integration of these photons within a certain length gives the intensity of the stimulus, which is a measure of perception. This implies that, perception is a probabilistic concept and therefore it is different than “seeing”, which is a goal oriented and therefore definitive. It is noteworthy to emphasize that the perception includes the brain processes to interpret an image of an object on the retina as existing object. That is, the image of an object on the retina cannot be taken granted for the realization of that object in the brain. Normally such a realization might most likely happen while at the same time it might not happen too depending on the circumstances although the latter is unlikely to occur. The brain processes are still not exactly known so that the ability to see an object without purposely searching for it is not a definitive process but a probabilistic process and we call this process as perception. This result clearly explains the relative importance of the front view as compared to side views in human vision.

B. Visual perception Model for an Orthogonal Geometry

In this part of this research we deal with visual perception in a more general geometry as shown in figure 5. For any point within that semi-enclosure, the perception can be theoretically computed. Without any restriction of generality, we consider the enclosure formed by three walls in the form of convex hull and by infinite vertical plane orthogonal to the paper plane. The wall dimensions are \( m_1, \ w_0 \) and \( m_2 \). We aim for to find the visual probability density with respect to the distance \( w \). For any point within the enclosure, we move the wall to the right so that the wall passes from the point in question. In such a case, the general shape of the geometry remains the same while only the geometric parameters \( m_1, \ m_2 \) and \( l_0 \) alters. This is indicated with the vertical broken line somewhere at the right hand side of the wall. In this geometry the whole region of interest is divided into four regions each of which is considered separately and the results are eventually combine for the final outcome. Regions I is formed by vertical plane and two broken lines denoted by \( x_1 \) and \( l_0 \). Region II is defined by the triangle formed by the broken line \( x_1 \) and two walls with the length \( m_1 \) and the upper partition of the leftmost wall which has two partitions with the length of \( w \) and \( w_0 \), as seen in figure 5. The regions II and IV are defined in a similar way at the lower side of the broken line with the length of \( l_0 \).

\[
\int_{-\infty}^{\infty} f_z(z) \, dz = \frac{l_0}{\pi} \int_{-\infty}^{\infty} \frac{1}{\sqrt{x^2 + z^2}} = 1
\]  

(14)
as it should be as pdf. The variation of \( f_z(z) \) is shown in figure 4.

Fig. 4. Perception pdf along the z axis parallel to the infinite plain in figure 1.
For region I, the visual probability density is found to be

\[ f_{1x}(x) = \frac{1}{\theta_1} \frac{l_o}{x\sqrt{x^2 - l_o^2}} \]  

(15)

for \(0 \leq \theta \leq \arctg \frac{w_o - w}{m_1}\).

(16)

where \(\theta_1\) is the normalization factor. As it is seen in (15), the probability density is dependent on the position along the leftmost wall side. To obtain the probability density of the position \(w\) along the wall side, we consider the joint probability distribution \(F(w,x)\) which is given by

\[ F(x,w) = \int_{l_o}^{b(w)} f_{1x}(x)dx = \int_{l_o}^{l_i/cos\theta} f_{1x}(x)dx \]  

(17)

where

\[ b = \frac{l_o}{\cos \theta} = l_o \frac{\sqrt{(w_o - w)^2}}{m_1} \]  

(18)

Substitution of (21) and (22) into (23) yields

\[ f_{1w}(w) = \frac{1}{\theta_1 m_1} \frac{1}{1 + \left(\frac{w_o - w}{m_1}\right)^2} \]  

(25)
This is the probability density of visual perception for the region I.

For the region II, we can write

\[ f_{II}(x) = \frac{1}{\phi_1} \frac{x}{\sqrt{x^2 - (w_o - w)^2}} \]  \hspace{1cm} (26)

for \( 0 \leq \phi_1 \leq \arctg \frac{w_o - w}{m_1} \).  \hspace{1cm} (27)

As it is seen in (26), the pdf is dependent on the position along the leftmost wall side. To obtain the pdf of the position w along the wall side with respect to the region II, we consider the joint probability distribution \( F_{II}(w,x) \) which is given by

\[ F_{II}(w,x) = \int_{w_o = w}^{w_o = w} f_{II}(x)dx \]  \hspace{1cm} (28)

where

\[ b(w) = \frac{w_o - w}{\cos \phi_1} = (w_o - w) \sqrt{1 + \left( \frac{m_1}{w_o - w} \right)^2} \]  \hspace{1cm} (29)

so that, the probability distribution \( F_{II}(w) \) is given by

\[ F_{II}(w) = \int_{\phi_1}^{\phi_2} f_{II}(x)dx \]  \hspace{1cm} (30)

The probability density of the variable w is obtained by differentiating the integral with respect to w. By the application of Leibniz rule to (20), we obtain

\[ f_w(w) = F(b(w),w) \frac{\partial b}{\partial w} \]  \hspace{1cm} (31)

where \( b(w) \) is given by (18), so that

\[ b(w) = \frac{w_o - w}{\cos \phi}_1 = (w_o - w) \sqrt{1 + \left( \frac{m_1}{w_o - w} \right)^2} \]  \hspace{1cm} (32)

\[ \frac{\partial b}{\partial w} = \frac{w_o - w}{m_1} \frac{m_1}{w_o - w} \sqrt{1 + \left( \frac{m_1}{w_o - w} \right)^2} \]  \hspace{1cm} (33)

Substitution of (32) and (33) into (31) yields

\[ f_{II}(w) = \frac{1}{\phi_1} \frac{m_1}{(w_o - w)^2} \frac{1}{1 + \left( \frac{m_1}{w_o - w} \right)^2} \]  \hspace{1cm} (34)

\[ = \frac{1}{\phi_1} \frac{m_1}{m_1^2 + (w_o - w)^2} \]  \hspace{1cm} (35)

The variation of \( F_I \) and \( F_{II} \) with w are given in figure 6

\[ F_I(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (36)

\[ F_{II}(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (37)

From (25) one obtains

\[ \theta_1 = m_1 \arctg \frac{w_o}{m_1} \]  \hspace{1cm} (38)

In the same way, from (34)

\[ F_{II}(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (39)

Since, for \( w = w_o \), \( F_{II}(w_o) = 1 \), from (39), it follows that

\[ \theta_1 = m_1 \arctg \frac{w_o}{m_1} \]  \hspace{1cm} (40)

In the same way, from (35)

\[ F_I(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (41)

Since, for \( w = w_o \), \( F_I(w_o) = 1 \), from (41), it follows that

\[ \theta_1 = m_1 \arctg \frac{w_o}{m_1} \]  \hspace{1cm} (42)

In the same way, from (37)

\[ F_I(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (43)

Since, for \( w = w_o \), \( F_I(w_o) = 1 \), from (43), it follows that

\[ \theta_1 = m_1 \arctg \frac{w_o}{m_1} \]  \hspace{1cm} (44)

In the same way, from (39)

\[ F_{II}(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (45)

Since, for \( w = w_o \), \( F_{II}(w_o) = 1 \), from (45), it follows that

\[ \theta_1 = m_1 \arctg \frac{w_o}{m_1} \]  \hspace{1cm} (46)

In the same way, from (41)

\[ F_I(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (47)

Since, for \( w = w_o \), \( F_I(w_o) = 1 \), from (47), it follows that

\[ \theta_1 = m_1 \arctg \frac{w_o}{m_1} \]  \hspace{1cm} (48)

In the same way, from (43)

\[ F_{II}(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (49)

Since, for \( w = w_o \), \( F_{II}(w_o) = 1 \), from (49), it follows that

\[ \theta_1 = m_1 \arctg \frac{w_o}{m_1} \]  \hspace{1cm} (50)

In the same way, from (45)

\[ F_I(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (51)

Since, for \( w = w_o \), \( F_I(w_o) = 1 \), from (51), it follows that

\[ \theta_1 = m_1 \arctg \frac{w_o}{m_1} \]  \hspace{1cm} (52)

In the same way, from (47)

\[ F_{II}(w) = \int_{\theta_1}^{\theta_2} \frac{m_j}{m_j^2 + (w_o - w)^2} dw \]  \hspace{1cm} (53)

Since, for \( w = w_o \), \( F_{II}(w_o) = 1 \), from (53), it follows that

\[ \theta_1 = m_1 \arctg \frac{w_o}{m_1} \]  \hspace{1cm} (54)
For $F_{III}(w)$ and $F_{IV}(w)$ using (35) and (36), we obtain

$$F_{III}(w) = \frac{1}{\theta_2} \int_0^w \frac{m_2}{m_2^2 + w^2} \, dw$$

$$= \frac{1}{\theta_2} m_2 \arctg \frac{w}{m_2}$$

Since, for $w=w_o$, $F_{III}(w_o)=1$, from (41), it follows that

$$\theta_2 = m_2 \arctg \frac{w}{m_2}$$

and

$$F_{IV}(w) = \frac{1}{\varphi_2} \int_0^w \frac{m_2}{m_2^2 + w^2} \, dw$$

$$= \frac{1}{\varphi_2} m_2 \arctg \frac{w}{m_2}$$

Since, for $w=w_o$, $F_{IV}(w_o)=1$, from (43), it follows that

$$\varphi_2 = m_2 \arctg \frac{w}{m_2}$$

For the total probability density

$$f_w(w) = \frac{\theta_1 f_{I}(w) + \varphi_1 f_{II}(w) + \theta_2 f_{III}(w) + \varphi_2 f_{IV}(w)}{\theta_1 + \varphi_1 + \theta_2 + \varphi_2}$$

$$= \frac{1}{\theta_1 + \varphi_1 + \theta_2 + \varphi_2} \left( \frac{m_1}{m_1^2 + (w_o - w)^2} + \frac{m_2}{m_2^2 + (w_o - w)^2} \right)$$

and finally the resulting probability density for the given geometry in figure 5, becomes

$$f_w(w) = \frac{1}{\theta_1 + \varphi_1 + \theta_2 + \varphi_2} \times \left( \frac{2m_1}{m_1^2 + (w_o - w)^2} + \frac{2m_2}{m_2^2 + w^2} \right)$$

The substitution of $\theta_1, \varphi_1, \theta_2, \varphi_2$ from (38), (40), (42), (44) into (46) yields the probability density as

$$f_w(w) = \frac{1}{m_1 \arctg \frac{w_o}{m_1} + m_2 \arctg \frac{w_o}{m_2}} \left( \frac{m_1}{m_1^2 + (w_o - w)^2} + \frac{m_2}{m_2^2 + w^2} \right)$$

The experimental setup in the virtual reality of the computer experiment reported in figure 7 is shown in figure 8 where the observation point is at the center of the wall along which the variable $w$ is defined.
The experiment similar to that shown in figure 8 is carried out in virtual reality with different experimental setting, namely \( m_1 = 2.5, m_2 = 1 \) and \( w_o = 5 \) that yields the probability density \( f_w(w) \) as shown in figure 11. The corresponding experimental setup is shown in figure 12. The determination of \( f_w(w) \) can be carried out at any place of the volume, which was considered in fig. 5. by simply passing a line parallel to the original line of observation. This new observation reject is shown with a broken line in figure 13. In this case only the parameters playing role on the probability density computation are changed, namely the length of the side walls \( m_1 \) and \( m_2 \). In other words new probability density formulation is not necessary only the same computer experiments are repeated with the new parameters.
Considering the variation of the probability density \( f_w(w) \) for the different geometries involved in the experiments it is interesting to note that for a geometry with a corridor like space, such as in the case of the experiment given in figure 9, the maximum of \( f_w(w) \) is at the middle of the line of observation that is at \( w_0/2 \). In rooms with wider proportions, such as in the case given in figure 7, two maxima occur between the extremities, \( w=0 \) and \( w=w_0 \) and the middle of the observation line. The wider the proportion of geometry, the closer to the extremities these maxima occur. Additionally we note that for asymmetrical spaces, that is geometries where \( m_1 \) and \( m_2 \) are different, the shape of \( f_w(w) \) becomes asymmetric and the higher values of the function occur near the extremity, which belongs to the longer wall. This result is striking and appears to be conforming common perceptual experience. By means of the computer experiments, in virtual reality the variations of \( f_w(w) \), which is shown in figures 7, 9, and 11 are verified.

IV. DISCUSSION AND CONCLUSION

The theory developed in this work which defines perception in probabilistic terms is verified by means of extensive computer experiments in virtual reality. From visual perception, other derivatives of perception can be obtained, like visual openness perception, visual privacy, visual color perception etc. In this respect, we have focused on visual openness perception where the change from visual to visual openness is accomplished via a mapping function and the work is reported in another publication [2]. Such perception related experiments have been carried out by means of virtual agents in the virtual reality. One example of such executions is shown in figure 11 where the perception of the agent is determined via the interacting rays simulating the vision process. Although, visual perception is commonly articulated in various contexts, generally it is used to convey a cognition related idea or message in a quite fuzzy form and this may be satisfactory in many instances. Such usage of perception is common in a daily life. However, in professional area, like architectural design or robotics, its demystification or precise description is necessary for the proficient executions. Since the perception concept is soft and thereby elusive there are certain difficulties to deal with it. For instance, how to quantify it or what are the parameters which play role on visual perception. Posit of this research is, that the perception process is very complex process including brain process. In fact, the latter, i.e., brain process about which what we know is highly limited, is final and therefore it is most important. Due to this complexity a probabilistic approach for a visual perception theory is very much appealing and the results obtained have direct implications which are in align with our common visual perception experiences, that we exercise every day. In this way, the work is a novel description of visual perception in probabilistic terms by clearly distinguishing between “seeing” and “perceiving”. Namely, “seeing” is a definitive process while visual perception is a probabilistic process.

This is because “seeing” is a goal oriented activity however perception is a matter of cognition or interpretation of information existing within the visual scope. This explains easily a common experience that human beings may overlook an object and search for it although such an overlook is not justified, and difficult to explain the phenomenon. It is firmly to conclude that visual perception is attached to the concept of distance the meaning is that in the formulation in some or other a distance is involved. Additionally, perception is to express in terms of intensity which is the integral of density. In the present context, these are the probability density and intensity of attention.

REFERENCES