Fusion of Perceptions for Perceptual Robotics

Ö. Ciftcioglu
Department of Building Technology
Delft University of Technology
Berlageweg 1, 2628 CR Delft, The Netherlands
o.ciftcioglu@tudelft.nl

M.S. Bittermann
Department of Building Technology
Delft University of Technology
Berlageweg 1, 2628 CR Delft, The Netherlands
m.s.bittermann@tudelft.nl

I.S. Sariyildiz
Department of Building Technology
Delft University of Technology
Berlageweg 1, 2628 CR Delft, The Netherlands
i.s.sariyildiz@tudelft.nl

***Abstract* - Fusion of perception information for perceptual robotics is described. The visual perception is mathematically modelled as a probabilistic process obtaining and interpreting visual data from an environment. The visual data is processed in a multiresolutional form via wavelet transform and optimally estimated via extended Kalman filtering in each resolution level and the outcomes are fused for each data block. The measurement involves visual perception in the virtual reality which has direct implications prominently in both design and perceptual robotics including navigation issues of actual autonomous robotics. For the interaction with the environment and visual data acquisition, the laser beams approach in robotics is considered and implemented by means of an agent in virtual reality which plays the role of robot in reality.**

I. Introduction

Visual perception is one of the important information source playing role on human's behavior. This behavior manifests itself clearly on many human activities. One such activity is design, in particular architectural design where the perception aspects in this domain are prominent to make optimal decisions about the form of the building or shaping its volumes for the achievement of high performance in actual use. The high performance includes effective functionality as well as high quality of life while the building is in actual use. Another activity comes into play when human-like behavior is to integrate with a human-like robot where robot is expected to mimic, to some extend, human behavior. The area is called perceptual robotics and in the last decade it received a growing interest for various applications spanning toys where especially emotions are expected to manifest and perceptual robotics. Perception measurements in architecture are another application for the purpose of design enhancement. By means of these examples, it is easy to realize that to integrate perceptual information into a machine-based system is a desirable achievement. Today, systems simulating vision, 3D scanners for instance, are quite well developed. However, the relation of perception to vision is not well described in mathematical terms in the literature. In this research, it is intended to shed some light on this relation and consequently better understanding the mechanism of perception for effective integration to interesting applications such as architectural design, perceptual robotics, as examples. The basic starting point is the nature of perception concept, which distinguishes itself from the concept of being able to "see". The seeing is a definitive process in contrast to the probabilistic concept of perception. The work is also intended to point out that, the overlapping areas of disciplines are growing in the modern era and the advanced methods of exact sciences, such as Kalman filtering and multiresolutinal decomposition can find applications in many diverse areas. To exemplify this, mention may be made of both *architecture and mathematics* and *perceptual robotics*. The organisation of the paper is as follows. Section two describes the multiresolutional filtering using wavelet transform. Section three gives the outcomes of experiments for fusion of perceptions investigating the merits of multiresolutional approach for perception fusion. This is followed by conclusions in section four.

II. Probabilistic Model of Visual Perception

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

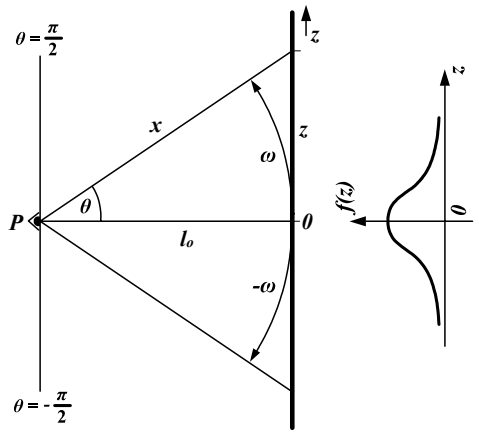

Fig. 1 The geometry of visual perception from a top view where P represents the position of eye, looking at a vertical plane with a distance l_o to the eye; $f_z(z)$ is the probability density function in z-direction

In figure 1, 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 θ. 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 and l. From figure 1, this is given by

$$x = \frac{l_o}{\cos(\theta)} \tag{1}$$

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 positing 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_θ is given by

$$f_\theta = \frac{1}{\pi/2} \tag{2}$$

Since θ 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* [1] and solve the equation

$$x = g(\theta) \tag{3}$$

for θ in terms of x. If $\theta_1, \theta_2, \ldots, \theta_n, \ldots$ are all its *real* roots,
$x = g(\theta_1) = g(\theta_2) = \ldots\ldots = g(\theta_n) = \ldots$.
then

$$f_x(x) = \frac{f_\theta(\theta_1)}{|g'(\theta_1)|} + \ldots + \frac{f_\theta(\theta_2)}{|g'(\theta_2)|} + \ldots + \frac{f_x(\theta_n)}{|g'(\theta_n)|} + \ldots \tag{4}$$

According the theorem above,

$$g'(\theta) = \frac{l_o \sin(\vartheta)}{\cos^2(\theta)} \tag{5}$$

Between $\theta = -\pi/4$ and $\theta = +\pi/4$,

$$g(\theta) = \frac{l_o}{\cos(\theta)} \tag{6}$$

has two roots, which are equal and given by

$$\theta_{1,2} = \arccos\left(\frac{l_o}{x}\right) \tag{7}$$

Using (7) in (5), we obtain

$$g'(\theta) = \frac{x\sqrt{x^2 - l_o^2}}{l_o} \tag{8}$$

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

$$f_x(x) = \frac{4}{\pi} \frac{l_o}{x\sqrt{x^2 - l_o^2}} \tag{9}$$

for the interval $l_o \le x \le \sqrt{2}\, l_o$. For this interval, the integration below becomes

$$\int_{l_o}^{\sqrt{2}\, l_o} f_x(x)\, dx = \frac{4}{\pi} \int_{l_o}^{\sqrt{2}\, l_o} \frac{l_o}{x\sqrt{x^2 - l_o^2}} = 1 \tag{10}$$

as it should be as pdf. The sketch of $f_x(x)$ vs x is given in figure 2. 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 $\theta \cong 0$, that is perpendicular direction to the plane. This striking result is in conformity with the common human experience as to visual perception.

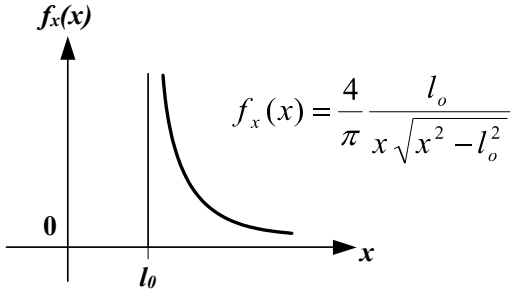

Fig. 2 Sketch explaining the relative importance of the viewing direction for visual perception

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.

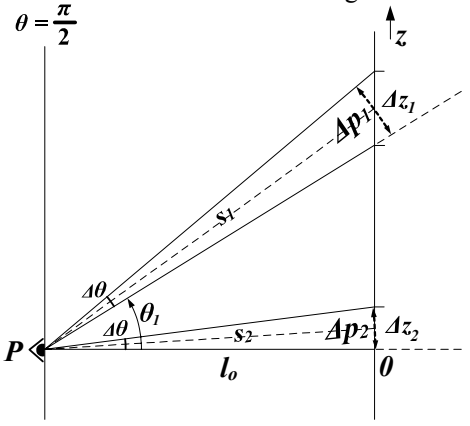

Fig. 3 Sketch explaining the relative importance of the viewing direction for visual perception.

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 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. The perception is associated with distance. This distance is designated as l_o in (9). From visual perception, one can obtain several visual perception derivatives, such as visual openness perception [2,3], visual privacy, visual accessibility etc.

For visual perception measurements, one can use a laser source at the location of robot vision system. The vision is simulated by sending random laser beams to the environment and receiving the backscattered beams afterwards. The probability density of such beams out of the laser source and received after backscattering is given by (9) as to the geometry in figure 1. These backscattered beams and the associated distances are recorded. These beams with respect to their backscattering distances are mapped to visual openness perception via a mapping function. In particular, in this research, this function is given by

$$y(t) = 1 - e^{-x(t)/\tau} \tag{11}$$

where x(t) represents the backscattering distance associated with beam. This is schematically shown in figure 4 with the associated probability density functions $f_x(x)$ and $f_y(y)$.

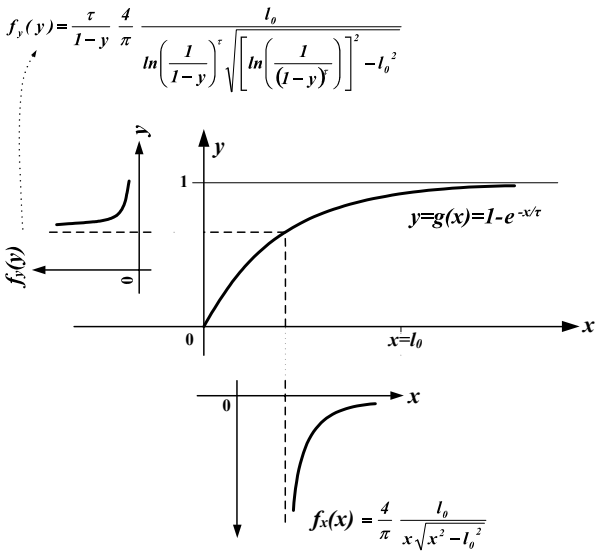

Fig. 4. Schematic representation of the probabilistic perception process via an exponential mapping(cognition) function y=f(x).

The actual implementation of this research is made in the virtual reality where a virtual agent plays the role of human robot. All perceptual robot behaviour is simulated in this environment to exercise the research outcome extensively without hardware or environmental limitations. However, the transfer of the perception technology being developed to the robotics is the final goal. A typical visual openness perception measurement with the virtual agent in real-time is shown in figure 5 where the vision beams underlying the measurements together with the plot of real-time measurement outcomes are clearly seen.

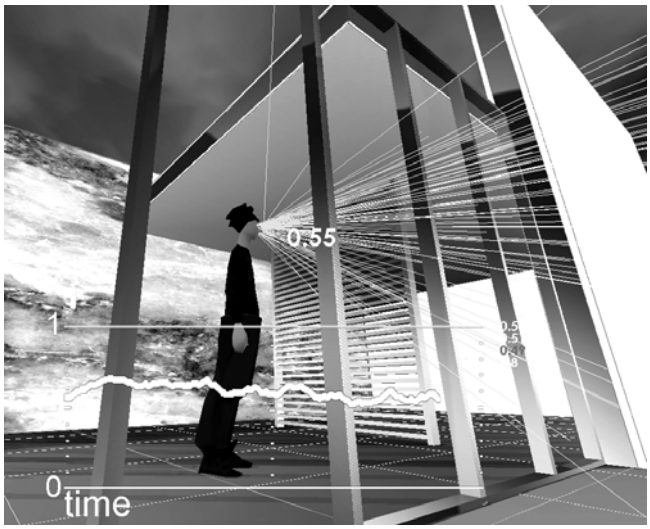

Fig. 5. Visual openness perception measurements via an agent in virtual reality playing the role of a robot in reality. The rays interact with the environment and provide the measurement data.

The visual openness perception is computed via exponential averaging [4] of the distances associated with the backscattered and mapped beams. A typical measurement is shown in figure 6 where the delay of the measurement due to the time constant (τ) of the exponential averaging is clearly seen.

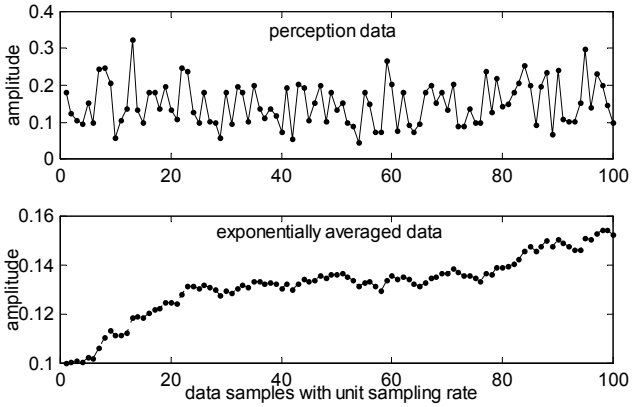

Fig. 6 The perception data (upper) and visual openness perception measurement outcome (lower) which is obtained by mapping the data for visual openness perception which is exponentially averaged with a time-constant τ for smoothing.

In real time perception measurements, the measurements should have minimum delay next to accurate determinations. This is accomplished by first decomposing the perception data into multiresolutional form by wavelet transform and at each resolution the extended Kalman filtering is applied for respective estimations. The estimated perceptions are fused to obtain the accurate perception determination at the highest resolution. At the same time, the delay due to time constant τ is eliminated and swift adaptation of the measurement system to the changing scenes in real time is achieved. In the measurement system, the mapping function plays important role since it contains the non-linear brain processes integrated to the backscattered data; that is the final interpretation of perception is made after the mapping function outcome which represents the final interpretation in human brain. It is important to note that, such a mapping function can be determined experimentally via experiential statements of a group of human in the framework of a cognition research. In this case, approximation of this curve by fuzzy logic has prominent importance where the placement of membership functions reflect the precise implementation of these research outcomes. This is exemplified in figure 7. The efficiency of fuzzy logic for stochastic modelling is discussed in another research [5].

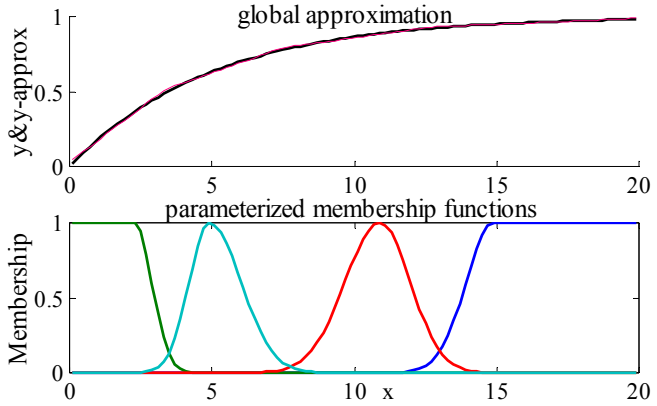

Fig. 7 Using four fuzzy sets, the fuzzy logic approximation of the mapping function representing the cognition process for perception in brain

II. Multiresolutional Filtering Using Wavelet Transform

In this section the multiresolutional filtering (MF), as proposed by Hong [6] will be briefly explained and presented. However, since Kalman filter underlies the MF algorithm firstly, a brief description of Kalman filter is presented.

A. The Kalman Filter

Consider a linear stochastic system to describe the propagation in time of a state vector X_t:

$$X_{t_k} = \mathbf{\Phi}(t_k, t_{k-1}) X_{t_{k-1}} + B(t_k) u_{t_k} + G(t_k) W_{t_k}, k = 1,2,.. \qquad (12)$$
$$X_{t_o} = X_o$$

where X_{t_k} is an n-vector state process, $\mathbf{\Phi}(t_k, t_{k-1})$ is n×n system dynamics matrix, $B(t_k)$ is an n×r input matrix , u_{t_k} ia an r-vector deterministic input, $G(t_k)$ ia an n×p noise input matrix and w_{t_k} is a p-vector white Gaussian noise process.

$$\{W_{t_k} W_{t_k}^T\} = Q(k) \qquad (13)$$

Measurements are available at times points t_1, t_2, ..and are modelled by

$$Z_{t_k} = \mathbf{C}(t_k) X(t_k) + V_{t_k} \qquad (14)$$

where Z_{t_k} is the c-vector measurement process, $\mathbf{C}(t_k)$ is c×n measurement matrix and V_{t_k} is an c-vector white Gaussian noise process with statistics

$$\{V_{t_k} V_{t_k}^T\} = R(k) \qquad (15)$$

The optimal state estimate is propagated from measurement time t_{k-1} to measurement time t_k by the equations

$$\begin{aligned} X(k \mid k-1) &= \mathbf{\Phi}(t_k, t_{k-1}) X(k-1 \mid k-1) + B(t_k) u_{t_k} \\ P(k \mid k-1) &= \mathbf{\Phi}(t_k, t_{k-1}) P(k-1 \mid k-1) \mathbf{\Phi}(t_k, t_{k-1})^T + \\ &\quad G(t_k) Q(k) G(t_k)^T \end{aligned} \qquad (16)$$

where P is the covariance matrix. At measurement time t_k, the measurement Z_{t_k} becomes available. The estimate is updated by the equation

$$\begin{aligned} X(k \mid k) &= X(k \mid k-1) + K(k)\left[Z_{t_k} - C(t_k) X(k \mid k-1)\right] \\ P(k \mid k) &= P(k \mid k-1) - K(k) C(t_k) P(k \mid k-1) \\ K(k) &= P(k \mid k-1) C(t_k)^T \left[C(t_k) P(k \mid k-1) M(t_k)^T + R(k)\right]^{-1} \end{aligned} \qquad (17)$$

K is the filter gain. Since (11) is a non-linear, linear model does not provide a valid description. Therefore, we consider a non-linear stochastic system

$$X_{t_k} = \mathbf{\Phi}(X_{t_k}, t_k, t_{k-1}) + B(t_k) u_{t_k} + G(t_k) W_{t_k}, k = 1,2,..$$
$$X_{t_o} = X_o \qquad (18)$$

where $\mathbf{\Phi}(X_{t_k}, t_k, t_{k-1})$ is an n-vector describing the system dynamics. The measurements are modelled by the non-linear equation

$$Z_{t_k} = \mathbf{T}(X_{t_k}, t_k) + V_{t_k} \qquad (19)$$

where $\mathbf{T}(X_{t_k}, t_k)$ is a vector describing the relation between the state and the measurements. For a reference trajectory x_{t_k} the state equation (18) can be linearized by Taylor expansion around the point X_{t_k}, t_k, t_{k-1}, so that it yields

$$X_{t_k} = \boldsymbol{\Psi}(X_{t_{k-1}}, t_k, t_{k-1}) X_{t_{k-1}} - \boldsymbol{\Psi}(x_{t_{k-1}}, t_k, t_{k-1}) x_{t_{k-1}} + \boldsymbol{\Phi}(x_{t_{k-1}}, t_k, t_{k-1}) + B(t_k) u_{t_k} + G(t_k) W_{t_k}, \quad (20)$$

$$Z_{t_k} = \mathbf{M}(x_{t_k}, t_k)\, X_{t_k} - \mathbf{M}(x_{t_k}, t_k) x_{t_k} + \mathrm{T}(x_{t_k}, t_{k)+V_{t_k}} \quad (21)$$

where

$$\left[\boldsymbol{\Psi}(x_{t_{k-1}, t_{k-1}, t_k})\right]_{ij} = \frac{\partial\left(\boldsymbol{\Phi}(x_{t_{k-1}}, t_{k-1}, t_k)\right)_i}{\partial(x_{t_k})_j} \quad (22)$$

$$\left[\mathbf{M}(x_{t_k}, t_k)\right]_{ij} = \frac{\left(\partial(\mathbf{M}(x_{t_k}, t_k)\right)_i}{\partial(x_{t_k})_j} \quad (23)$$

and the approprimate linear observation equation

$$Z_{t_k} = \mathbf{M}(x_{t_k}, t_k) X_{t_k} - \mathbf{M}(x_{t_k}, t_k) x_{t_k} + \mathbf{T}(x_{t_k}, t_k) + V_{t_k} \quad (24)$$

Given the linearized model described, the standard Kalman filter is used to obtain the estimate of the state X_{t_k} and it covariance matrix. For the reference trajectory, the obvious choice is

$$x_{t_k} = \boldsymbol{\Phi}(x_{t_{k-1}}, t_{k-1}, t_k) + B(t_k) u_{t_k} \quad (25)$$

so that the reference trajectory is completely determined by the prior estimate of the state. This estimator is called the linearized Kalman filter. More information about the Kalman filter, can be found in [7-11].

B. The Wavelet Transform

The functions given by discrete data of the form $f(x_i)=d_i$, i=1,2,...m , can be represented in multiresolutional form in a dyadic structure as a counterpart of the continuous wavelet transformed time-frequency representation. The multiresolution theory can be conveniently described by the theory of function spaces. A function space is made of embedded subspaces V_m the limit of their union is $\mathbf{L}^2(\mathbf{R})$ where for each function f(x) $\in V_m$ we can write that

$$f(x) \in V_m \Leftrightarrow f(2x) \in V_{m-1}$$

$$.... \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset .....$$

In $\mathbf{L}^2(\mathbf{R})$, the functions

$$\phi_{m,n}(x) = 2^{-m/2} \phi(2^{-m} x - n) \quad (26)$$

form an orthonormal basis for V_m. These are called scaling functions and for m=0, we basically write

$$\phi_{0,n}(x) = \phi(x - n) \quad (27)$$

The function f(x) in each subspace can be expressed by these orthogonal base functions as approximation in such a way that $f_m(x) \in V_m$ and

$$f(x) = \lim_{m \to \infty} f_m(x) \quad (28)$$

All functions in V_m can be represented by using linear combinations of the scaling functions. In other words, $f_m(x)$ is an orthogonal projection of f(x) onto V_m,

$$f_m(x) = \sum_n < \phi_{m,n}(x), f(x) > \phi_{m,n}(x)$$

$$= \sum_n c_{m,n}\, \phi_{m,n}(x) \quad (29)$$

where $< \phi_{m,n}(x), f(x) >$ is the inner product:

$$< \phi_{m,n}(x), f(x) > = \int_{-\infty}^{\infty} \phi_{m,n}(x)\, f(x)\, dx \quad (30)$$

The difference spaces can be represented by W_m and are defined as the orthogonal complement of the spaces V_m with respect to V_{m-1},

$$V_{m-1} = V_m \oplus W_m$$

where V_m and W_m are orthogonal to each other. Now, let $\psi(x) = \psi_{0,0}(x)$ be a basis function of W_o. Note that $\psi_{0,0}(x) \in W_0 \subset V_{-1}$ and therefore can expressed in terms of basis functions $\phi_{-1,n}(x)$ and therefore, we can also define functions $\psi_{m,n}(x)$ that are shifted and dilated versions of one prototype function ψ(x) of the form

$$\psi_{m,n}(x) = 2^{-m/2} \psi(2^{-m} x - n) \quad (31)$$

The functions $\psi_{m,n}(x)$ are identical to the wavelets described before, after the discretization. There are strong relations between $\phi(x)$ and $\psi(x)$. The introduction of the wavelet functions enables us to write any function $f(x)$ in $\mathbf{L}^2(\mathbf{R})$ as a sum of projections on W_j, j∈ R, of the form

$$f(x) = \sum_{j=-\infty}^{\infty} w_j(x) \quad (32)$$

where

$$w_j(x) = \sum_k < \psi_{j,k}(x), f(x) > \Psi_{j,k}(x) \quad (33)$$

Considering a certain scale m, the function f(x) can be written as the sum of a low resolution part $f_m(x) \in V_m$ and the detail part which is constituted by the wavelets $w_j(x) \in W_j$ so that

$$f(x) = f_m(x) + \sum_{j=-\infty}^{m} w_j(x)$$

$$= \sum_n < \phi_{m,n}(x), f(x) > \phi_{m,n}(x) +$$

$$\sum_{j=-\infty}^{k}\sum_{k}^{m} < \psi_{j,k}(x), f(x) > \psi_{j,k}(x) \qquad (34)$$

which can be expressed as

$$f(x) \quad = \sum_{n} c_{m,n}\phi_{m,n}(x) + \sum_{j=-\infty}^{m}\sum_{k} d_{j,k}\psi_{j,k}(x) \qquad (35)$$

Above, the coefficients $d_{j,k}$ are known as the *wavelet coefficients*. From the preceding equation multiresolution decomposition is represented by an approximation i.e., the first term with $\phi_{m,n}(x)$ functions, and the detail part i.e., the second term with the $\psi_{j,k}(x)$ functions. The variable m indicates the scale and is called *scale factor* or *scale level*. If the scale level m is high, it indicates that the function in V_m is a coarse approximation of f(x), so the details are neglected. On the contrary, if the scale level is low, a detailed approximation of $f(x)$ is achieved. More information about the wavelet transform, can be found in [12-14].

C. Signal Decomposition and Reconstruction Using Haar Wavelets

The time series signal first decomposed to lower resolution levels using Haar wavelet. Haar wavelet is a two-tap high pass filter given by

$$g_{haar} = [g_1 \;\; g_2] = \left[\frac{1}{2}\sqrt{2} \quad -\frac{1}{2}\sqrt{2}\right] \qquad (36)$$

The two-tap Haar low pass filter coefficients are

$$h_{haar} = [h_1 \;\; h_2] = \left[\frac{1}{2}\sqrt{2} \quad \frac{1}{2}\sqrt{2}\right] \qquad (37)$$

The signals at the lower levels constitute the respective measurements and at each level extended Kalman filter is applied. Note that, these are *calculated* measurements and they contain less information than the original measurements. However, they can better capture certain information at lower resolutions as result of low-pass filtering during decomposition. The measurements at different resolution levels is shown in figure 8 and the decomposition of state variables at higher resolution level to lower resolution levels is shown in figure 9.

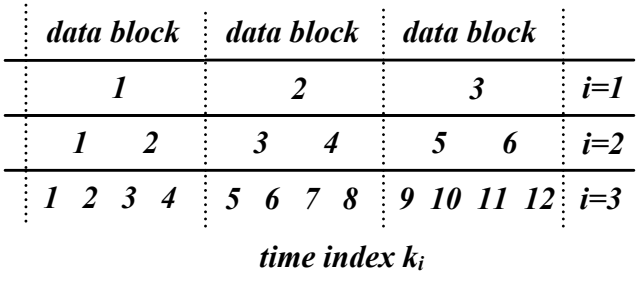

Fig. 8 Measurements at different resolution levels i for i=1,2,3

Note that, the update of the states is executed when a data block is ready. In this research, the number of resolution levels is N=3, and each data block contains 4 samples at the highest resolution level. The basic scheme for dynamic multiresolutional filtering is shown in figure 10.

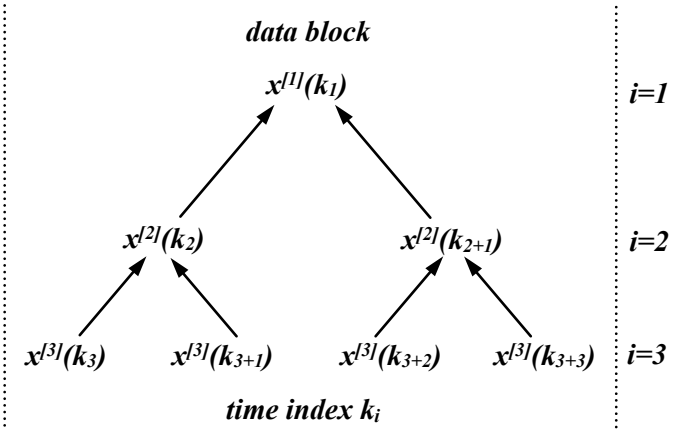

Fig. 9 Decomposition of the state variables at highest resolution level i=3 to lower resolution levels, i=1,2.

$X_{m+1|m}^{[N]}$ *updatewith* $Z^{[N]}$ $X_{m+1|m+1}^{[N]}$

$X_{m+1|m}^{[N-1]}$ *updatewith* $Z^{[N-1]}$ $X_{m+1|m+1}^{[N-1]}$

. . . . .

$X_{m+1|m}^{[1]}$ *updatewith* $Z^{[1]}$ $X_{m+1|m+1}^{[1]}$

data block n

$\Rightarrow$ *datablock* $n+1$

Fig. 10 Multiresolutinal decomposition during filtering

Once the N sequences of updated state variables end error covariances $X_{m+1|m+1}^{[N,i]}$ and $P_{m+1|m+1}^{[N,i]}$ for I=1,2,..,N, are determined, they must be fused to generate an optimal $X_{m+1|m+1}^{[NF]}$ and $X_{m+1|m+1}^{[NF]}$. For the minimum fusion error covariance $X_{m+1|m+1}^{[NF]}$, the fused estimate $X_{m+1|m+1}^{[NF]}$ is calculated as

$$X_{m+1|m+1}^{[NF]} = P_{m+1|m+1}^{[NF]}\left\{\left[\sum_{i=1}^{N}\left(P_{m+1|m+1}^{[N,i]}\right)X_{m+1|m}^{[N,i]} - \right]\right. \qquad (38$$

$$\left.(N-1)\left(P_{m+1|m}^{[N]}\right)^{-1}X_{m+1|m}^{[N]}\right)$$

where the minimum fusion error covariance $P_{m+1|m+1}^{[NF]}$ becomes

$$\left(P_{m+1|m+1}^{[NF]}\right)^{-1} = \sum_{i=1}^{N}\left(P_{m+1|m+1}^{[N,i]}\right)^{-1} - (N-1)\left(P_{m+1|m}^{[N]}\right)^{-1}. \quad (39)$$

The fused estimate $X_{m+1|m+1}^{[NF]}$ is a weighted summation of both predicted $X_{m+1|m}^{[N]}$ and updated $X_{m+1|m+1}^{[N,i]}$, for I=1,2,..,N. The sum of the weight factors equal to the identity I. This can be seen by substitution of $P_{m+1|m+1}^{[NF]}$ given above into the expression of $X_{m+1|m+1}^{[NF]}$ in (39).

III. Experiments for Fusion of Perceptions

The perception data subject to decomposition and information fusion is obtained by means a visual agent in the virtual reality environment. The rays stemming from the agent's eye interact with the environment and return back as result of backscattering. The distance associated with these rays is used in the mapping function to estimate the perception as depicted in figure 4. The average number of backscattered rays is the measure of perception in the form of intensity. This intensity can be calculated by means of integration of the associated probability density. The exponential averaging described in the preceding section delivers the average value of this intensity. Such a system is established for real-time visual openness perception measurements. The perception data is a set of random samples and it is colored data due to correlations peculiar to the space subjected to measurement. The experiments presented below are for 100 samples. In the virtual reality environment the frame rate is about 20 frames/sec so that the experiments take approximately 5 seconds. For a stationary viewing position, the perception data, its Kalman filtered counterpart and the input signal to the system is shown in figure 11. Apparently, the stationary input provides a stationary output.

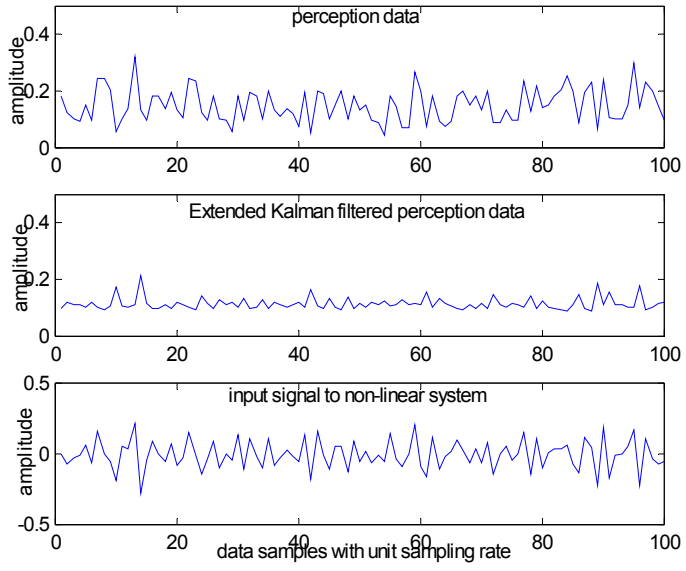

Fig. 11 Visual openness measurement for a stationary viewing position.

For the case the measurements are not stationary while the scene is changing during the constant movement of the agent, a varying measurement outcome is shown in figure 12. There are three lines plotted in this figure. The solid line is the Kalman filtering estimation at the highest resolution of the perception measurement data. The cross symbols connecting the lines represent the measurement data set. The outcome of the multiresolutional fusion process is given with the dot-dashed line. The upper plot in figure 12 is zoomed and presented in the lower plot for explicit illustration of the experimental outcomes. The same figure as figure 12 with a different zooming range and the zooming power is given in figure 13.

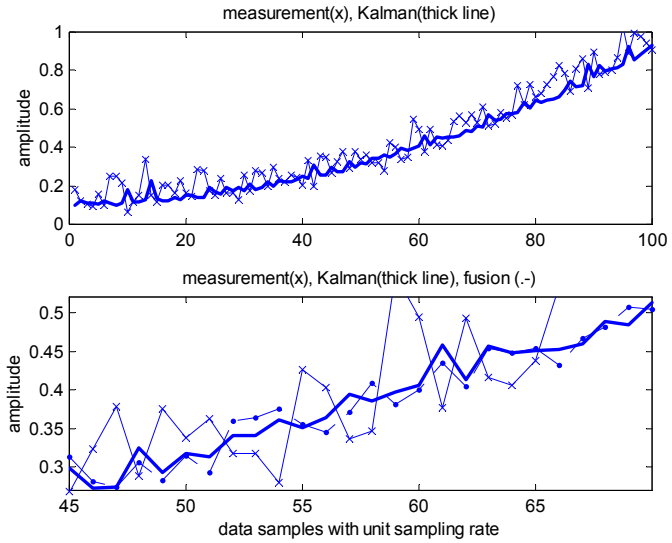

Fig. 12 Visual openness measurement outcome from a moving virtual agent

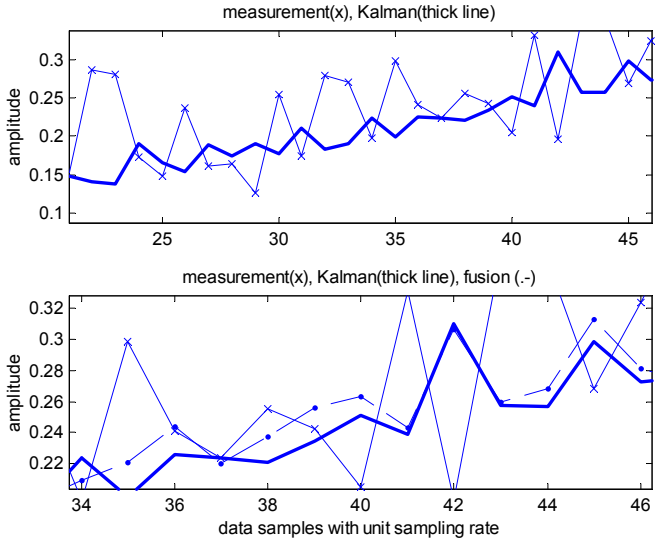

Fig. 13 Visual openness measurement outcome from a moving virtual agent

From the experiments it is seen that, the Kalman filtering is effective to model the noise of the perception signal and give accurate estimation of the perception measurement. From a

stationary measurement position the estimated perception is also stationary and the statistical variations are minimized. For a non-stationary observation, the scene is not stationary and the perception measurements are subject to varying. The filtered perception measurements provide the optimal estimation of the visual perception where the estimation error is minimized so that the noise on the measurement data greatly alleviated. At the same time swift adaptation is obtained with the result that the delay in adaptation occurring in exponential averaging is eliminated. It is interesting to note that the multiresolutional fusion outcomes do not present any significant improved estimation as compared with the Kalman filtering estimation at the highest resolution level, in this case. However, a slight difference between these two estimates is noticeable in favor of the multiresolutional case. It is noteworthy to mention that, the multiresolutional approach presented here uses *calculated* measurements in the lower resolutions. Therefore, since the information content is the same in both cases, a non-significant difference in these results is not surprising. However, the multiresolutional fusion is still important alternative for improved estimation since it is possible to use different sensors at each resolution level and to obtain independent information subject to fusion. In the virtual environment, this means different independent virtual agents at each level and this can easily be added to the present measurement system to improve the measurement system performance.

IV. Conclusions

Fusion of perceptions is investigated for perception measurements where accurate estimations are aimed. For this aim, several measurements in different resolutional levels are considered where each outcome at each level is combined with the others and a final outcome is obtained. This is commonly referred to as data/sensor fusion. In this research, measurements are the perception of human where human perception is modelled with probabilistic considerations so that the measurements are in the form of random data. For accurate estimations using the measurement samples, optimal filtering, namely extended Kalman filtering is applied at each multiresolutional level. The multiresolutional sensor fusion outcomes are compared with the Kalman filtering outcomes at the highest resolution level. The difference between the outcomes is found to be noticeable but not significant. This is attributed to "calculated" sensors rather than independent sensors in the lower resolutions. In the present research, the virtual agent provides the measurement data so that in the multiresolutional case to increase the number of agents is another alternative for accurate perception measurements. Next to optimality of Kalman filtering for estimation of perception in a fixed observation location, it is fast enough to follow the perception variations of a moving agent with changing scenes, maintaining the same performance. Following the theoretical considerations developed in this research, the present experiments are carried out in the virtual reality environment, in real-time. However, the same executions can be made in real life environment exercised by an autonomous robot. Therefore, the implication of this research extends to autonomous robotics as well as perceptual robotics. At the same time from the design viewpoint, visual perception is an important concept in building design in architecture and quantification of visual perception in the form of measurements is an important step for design enhancement. It is noteworthy to mention that, as the scientific disciplines are getting more and more overlapping in their field of interests due to due increasing need of interdisciplinary cooperation, the present research is an exemplary endeavour to integrate advanced exact science methodologies into design environment that refers to architectural design as well as engineering design.

References

[1] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, 1965.

[2] Ö. Ciftcioglu, M. Bittermann and I.S. Sariyildiz, "Studies on visual perception for perceptual robotics", Proc. *ICINCO 2006 3rd Int. Conference on Informatics in Control, Automation and Robotics*, August 1 - 5, 2006, Setubal, Portugal.

[3] M. Bittermann and Ö. Ciftcioglu, "Real-time measurement of perceptual qualities in conceptual design", Proc. *International Symposium series on Tools and Methods of Competitive Engineering*, *TMCE 2006*, April 18–22, 2006, Ljubljana, Slovenia.

[4] T.T.J.M. Peeters and Ö. Ciftcioglu, "Statistics on exponential averaging of periodigrams," in *IEEE Trans. on Signal Processing*,vol.43, no.7, 1995, pp. 1631-1636.

[5] Ö. Ciftcioglu, "On the efficiency of fuzzy logic for stochastic modeling", *NAFIPS'06,* Montreal, Concordia University, June 3-6, 2006, Concordia University, Montreal, Quebec, Canada .

[6] L. Hong, "Multiresolutional filtering using wavelet transform", *IEEE Transactions on Aerospace and Electronic Systems*, vol.29, no.4, pp.1244-1251, 1993.

[7] A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York 1970.

[8] P.S. Maybeck, *Stochastic Models, Estimation and Control*, Vol I, Academic Press, New York, 1979.

[9] P.S. Maybeck, *Stochastic Models, Estimation and Control*, Vol II, Academic Press, New York, 1982.

[10] B.D.O. Anderson and J.B. Moore, *Optimal Filtering*, Prectice-Hall, Englewood Cliffs, New Jersey, 1979.

[11] R.G. Brown, *Introduction to Random Signal Analysis and Kalman Filtering*, John Wiley & Sons, New York 1983.

[12] S. Mallat, *A Wavelet Tour of Signal Processing,* Associated Press, New York, 1999.

[13] S.G. Mallat, "A theory for multiresolution signal decomposition:the wavelet representation", *IEEE trans. Pattern Analysis and Machine Intelligence*, vol.11, isuue 7, pp.674-693, July 1989.

[14] D.B. Percival and A.T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge Univ. Press, 2000.