Sensor Data Fusion in Autonomous Robotics

Ö. Ciftcioglu, M. S. Bittermann and I. S. Sariyildiz
Faculty of Architecture, Delft University of Technology, The Netherlands
o.ciftcioglu@tudelft.nl

Abstract

Studies on sensor data fusion in autonomous perceptual robotics are described. The visual perception is represented by a probabilistic model, where the model receives and interprets visual data from the environment in real-time. The perception obtained in the form of measurements in 2D is used for perceptual robot navigation. By means of this twofold gain is obtained; while the autonomous robot is navigated, it is equipped with some human-like behaviour. 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 improved estimation of the trajectory. Various forms of sensor-data fusion is described. The perceptual robotics experiments are carried out in virtual reality for the demonstration of the feasibility of the investigations in this domain. The improvement on the trajectory estimation by means of sensor/data fusion is demonstrated.

1. Introduction

Visual perception is one of the important information sources playing role on human's behavior A research is presented earlier which demystifies the concepts of perception and attention as to vision from their verbal description to a scientific formulation [1]. For the verification of the theoretical study done in that work, outcomes from the model is presently implemented in an avatar-robot in virtual reality. The perceptual approach for autonomous movement in robotics is also important as the perception is very appropriate in a dynamic environment, where predefined trajectory or trajectory conditions like occasional obstacles or hindrances are duly taken care of. On the other hand, the approach can better deal with the complexity of environments by processing environmental information selectively. To demonstrate this, a trajectory is defined and the same trajectory is estimated from the simulated perception measurements by means of sensor fusion which involves Kalman filtering and wavelet transform for the multiresolutional representation of the sensory data. The paper gives some detailed information about the sensor data fusion accomplished and experimental results in the virtual reality. The organization of the paper is as follows. Section two gives the brief description of the perception model developed in the framework of ongoing perceptual robotics research. Section three is concerned with multiresolutional filtering via wavelets and Kalman for sensor fusion yielding enhanced trajectory estimation. Section four is reserved for the experimental studies. Section five gives some discussions and the conclusions.

2. Probabilistic model of visual perception

We start with the basics of the perception process with a simple and special, yet fundamental orthogonal visual geometry. It is shown in Fig. 1.

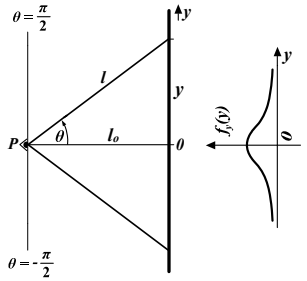

Fig. 1 Orthogonal geometry of perception.

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 visual perception is associated with distance, it is

straightforward to proceed to express the distance of visual perception x in terms of θ and l_o. From figure 1, this is given by

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

Since we consider that 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 $\theta = \pm \pi/2$, 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 $f_l(l)$ of this random variable is computed as [44]

$$f_l(l) = \frac{2}{\pi} \frac{l_o}{l\sqrt{l^2 - l_o^2}} \tag{3}$$

for the interval $l_o \le x \le \infty$. Considering that

$$tg(\theta) = \frac{y}{l_o} \tag{4}$$

and by means of pdf calculation similar to that to obtain $f_x(x)$ one can obtain $f_y(y)$ as [1].

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

for the interval $-\infty \le y \le \infty$. (9) and (11) are dual representation of the same phenomenon. The probability density functions $f_l(l)$ and $f_y(y)$ are defined as *attention* in the terminology of cognition.

By the help of the results given by (3) and (5) two essential applications in design and robotics are described in a previous research [2]. The fundamental orthogonal visual geometry is extended to a general visual geometry to explore the further properties of the perception phenomenon [3]. In this geometry the earlier special geometry the orthogonality condition of the infinite plane is relaxed. This geometry is shown in Fig. 2 where the attentions at the points O and O' are subject to computation, with the same axiomatic foundation of the probabilistic theory, as before. Since the geometry is symmetrical with respect to x axis, we consider only the upper domain of the axis without loss of the generality. The probability density with respect to the variable r is given by

$$f_r^*(r) = \frac{l_o}{\pi/2} \frac{1}{r^2 - 2l_o r\cos\varphi + l_o^2} \tag{6}$$

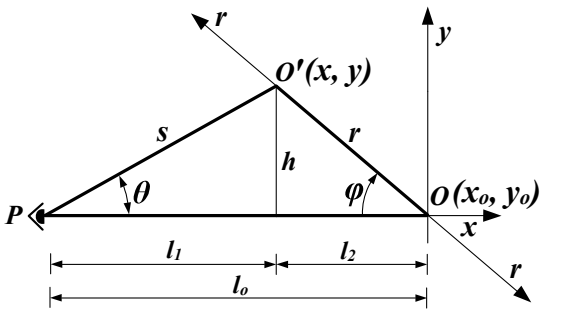

Fig.2 General geometry of visual perception.

The pdf has several interesting features. First, for $\varphi = \pi/2$, it boils down

$$f_y(r) = \frac{l_o}{\pi/2} \frac{1}{\left(r^2 + l_o^2\right)}$$

An interesting point is that when $\varphi \to 0$ but $r \neq 0$. This means O' is on the gaze line from P to O. For the case O' is between P and O, $f_r(r)$ becomes

$$f_r(r) = \frac{l_o}{\pi} \frac{1}{(l_o \pm r)^2} \tag{7}$$

In (7) for $r \to l_0$ $f_r(r) \to \infty$. This case is similar to that in (3) where $l \to l_0$ $f_l(l) \to \infty$. The actual $f_r(r)$ is obtained as the intersection of a vertical plane passing from the origin O and the surface. The analytical expression of this intersection is given by (6). The 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 robotic perception measurement with the virtual agent in real-time is shown in Fig. 3 where the vision beams underlying the measurements together with the plot of real-time measurement outcomes are clearly seen. The intensity of the visual beams of the virtual agent is in a Gaussian-like form as seen in figure 1. The exact form is given by (5), as a probability density. The visual perception is computed via exponential averaging of the distances associated with the backscattered beams. The rays stemming from the agent's eye interact with the environment. The interaction points with the environment are recorded and the position is identified as the exponentially averaged value of the coordinates. This means there is some delay in the measurements, as delay of perception, depending on the time constant involved in the exponential averaging. The autonomous movement of the agent or avatar/virtual-robot is accomplished by multiresolutional filtering which compensates for this delay.

3. Multiresolutional filtering using Kalman filtering and the wavelet transform

3.1. The Kalman Filter

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

$$X_{t_k} = \Phi(t_k, t_{k-1})X_{t_{k-1}} + B(t_k)u_{t_k} + G(t_k)W_{t_k}, k = 1,2,.. \qquad (8)$$
$$X_{t_o} = X_o$$

where X_{t_k} is an n-vector state process, $\Phi(t_k, t_{k-1})$ 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 $n \times p$ noise input matrix and w_{t_k} is a p-vector white Gaussian noise process.

$$\{W_{t_k} W_{t_k}^T\} = Q(k) \qquad (9)$$

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

$$Z_{t_k} = \mathbf{C}(t_k) X(t_k) + V_{t_k} \qquad (10)$$

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

$$\{V_{t_k} V_{t_k}^T\} = R(k) \qquad (11)$$

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

$$X(k \mid k-1) = \Phi(t_k, t_{k-1})X(k-1 \mid k-1) + B(t_k)u_{t_k}$$
$$P(k \mid k-1) = \Phi(t_k, t_{k-1})P(k-1 \mid k-1)\Phi(t_k, t_{k-1})^T + G(t_k)Q(k)G(t_k)^T \qquad (12)$$

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

$$X(k \mid k) = X(k \mid k-1) + K(k)\left[Z_{t_k} - C(t_k)X(k \mid k-1)\right]$$
$$P(k \mid k) = P(k \mid k-1) - K(k)C(t_k)P(k \mid k-1) \qquad (13)$$
$$K(k) = P(k \mid k-1)C(t_k)^T\left[C(t_k)P(k \mid k-1)M(t_k)^T + R(k)\right]^{-1}$$

K is the filter gain. Since the avatar's movement trajectory is non-linear, a linear model does not provide a valid description. Therefore, we consider a non-linear stochastic system

$$X_{t_k} = \Phi(x_{t_k}, t_k, t_{k-1}) + B(t_k)u_{t_k} + G(t_k)W_{t_k}, k = 1,2,.. \qquad (14)$$
$$X_{t_o} = X_o$$

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

$$Z_{t_k} = \boldsymbol{c}(X_{t_k}, t_k) + V_{t_k} \qquad (15)$$

where $\boldsymbol{c}(X_{t_k}, t_k)$ is a vector describing the relation between the state and the measurements. For a reference trajectory x_{t_k} the state equation (14) can be linearized by Taylor expansion around the point X_{t_k}, t_k, t_{k-1}, so that it yields

$$X_{t_k} = \Psi(X_{t_{k-1}}, t_k, t_{k-1})X_{t_{k-1}} - \Psi(x_{t_{k-1}}, t_k, t_{k-1})x_{t_{k-1}} + \Phi(x_{t_{k-1}}, t_k, t_{k-1}) + B(t_k)u_{t_k} + G(t_k)W_{t_k}, \qquad (16)$$

$$Z_{t_k} = \mathbf{M}(x_{t_k}, t_k)X_{t_k} - \mathbf{M}(x_{t_k}, t_k)x_{t_k} + \boldsymbol{c}(x_{t_k}, t_{k)+V_{t_k}} \qquad (17)$$

where

$$\left[\Psi(x_{t_{k-1}, t_{k-1}, t_k})\right]_{ij} = \frac{\partial\left(\Phi(x_{t_{k-1}}, t_{k-1}, t_k)\right)_i}{\partial(x_{t_k})_j} \qquad (18)$$

$$\left[\mathbf{M}(x_{t_k}, t_k)\right]_{ij} = \frac{\left(\partial(\boldsymbol{c}(x_{t_k}, t_k)\right)_i}{\partial(x_{t_k})_j} \qquad (19)$$

and the approprimate linear observation equation

$$Z_{t_k} = \mathbf{M}(x_{t_k}, t_k)X_{t_k} - \mathbf{M}(x_{t_k}, t_k)x_{t_k} + \boldsymbol{c}(x_{t_k}, t_k) + V_{t_k} \qquad (20)$$

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

$$x_{t_k} = \Phi(x_{t_{k-1}}, t_{k-1}, t_k) + B(t_k)u_{t_k}; \quad x_{t_o} = X_o \qquad (21)$$

so that the reference trajectory is completely determined by the prior estimate of the state. This estimator is called the linearized/extended Kalman filter (EKF). More information about the Kalman filter, can be found in [4-6].

3.2. Decomposition and reconstruction using Haar wavelets

In this section the multiresolutional filtering (MF), as proposed by Hong [7] will be briefly explained and presented. The decomposition of state variables to lower resolution levels and reconstruction from the lower resolution levels, is accomplished using discrete wavelet transformation and inverse wavelet transformation [8] via quadrature mirror filters (QMF). Two-tap Haar low-pass (h_{haar}) and high pass (g_{haar}) filters are given by

$$h_{haar} = [h_1 \; h_2] = \left[\frac{1}{2}\sqrt{2} \quad \frac{1}{2}\sqrt{2}\right] \qquad (23)$$

$$g_{haar} = [g_1 \; g_2] = \left[\frac{1}{2}\sqrt{2} \quad -\frac{1}{2}\sqrt{2}\right] \qquad (24)$$

Within a block, the data at each level represents the respective measurements and at each level extended Kalman filter is applied. The data blocks for each set of measurements are shown in Fig. 6. Note that, in this scheme 4-step ahead estimation is performed due to

blockwise Kalman filtering estimation. However a more effective estimation scheme used in this work is shown in Fig. 7 where the estimation in a block is updated based on the last measurement of the previous block.

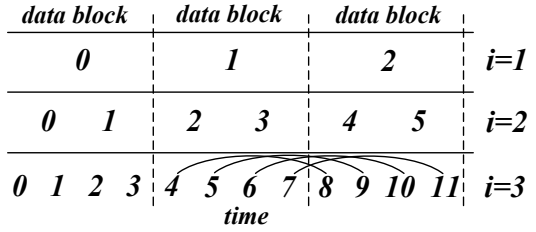

Fig. 6. Measurements at different resolution levels i for i=1,2,3.

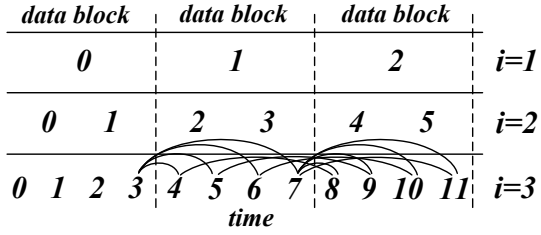

Fig. 7 Measurements at different resolution levels i for i=1,2,3.

The basic update scheme for dynamic multiresolutional filtering is shown in Fig. 8 where at each resolutional level, when the measurement is available, the state variables are updated and when the block is complete the inverse wavelet transform and fusion is performed. Note that, during the inverse transformation wavelet coefficients $Y_{m+1|m+1}^{[i]}$ saved aside are used.

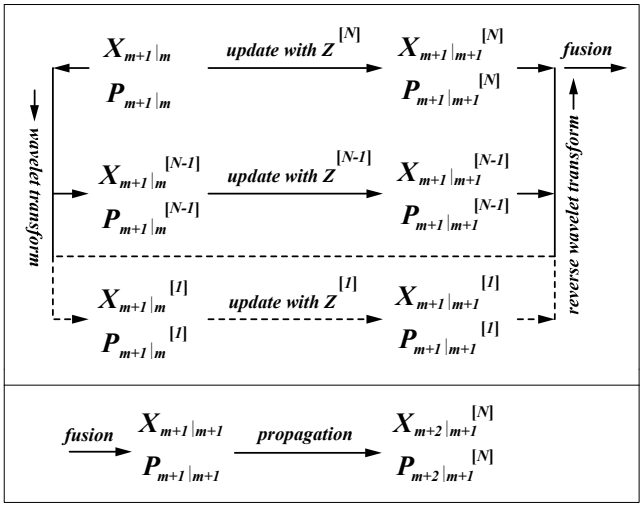

Fig. 8 Multiresolutinal decomposition during filtering

Explicitly

$$P_{m+1|m}^{[i]} = \begin{bmatrix} P_{XXm+1|m}^{[i]} & P_{XYm+1|m}^{[i]} \\ P_{YXm+1|m}^{[i]} & P_{YYm+1|m}^{[i]} \end{bmatrix} \quad (25)$$

$$X_{m+1|m+1}^{[i]} = X_{m+1|m}^{[i]} + K_{m+1}^{[i]}(Z_{m+1}^{[i]} - C_{m+1}^{[i]} X_{m+1|m}^{[i]} \quad (26)$$

and

$$P_{XXm+1|m+1}^{[i]} = (I - K_{m+1}^{[i]} C_{m+1}^{[i]}) P_{XXm+1|m}^{[i]} \quad (27)$$

As $X_{m+1|m}^{[i]}$ *and* $Y_{m+1|m}^{[i]}$, $Y_{m+1|m}^{[i+1]}$ are correlated, the covariance matrices $P_{XYm+1|m}^{[i]}$ and $P_{YXm+1|m}^{[i]}$ are also updated. However. the wavelet coefficients $Y_{m+1|m}^{[i]}$ and their covariance matrices $P_{YYm+1|m}^{[i]}$ however are not updated. The minimum variance Kalman gain matrix $K_{m+1}^{[i]}$ at each level, is determined by

$$K_{m+1}^{[i]} = P_{XXm+1|m}^{[i]} C_{m+1}^{[i]T} \left(C_{m+1}^{[i]} P_{XXm+1|m}^{[i]} C_{m+1}^{[i]T} + R_{m+1}^{[i]} \right)^{-1} (28)$$

Once, within the moving window, the sequences of updated state variables and 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 $X_{m+1|m+1}^{[NF]}$. $X_{m+1|m+1}^{[NF]}$ 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

$$X_{m+1|m+1}^{[NF]} = P_{m+1|m+1}^{[NF]} \left[\sum_{i=1}^{N} \left(P_{m+1|m+1}^{[N,i]} \right) X_{m+1|m}^{[N,i]} - \right] \quad (29)$$
$$(N-1) \left(P_{m+1|m}^{[N]} \right)^{-1} X_{m+1|m}^{[N]}$$

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

$$\left(P_{m+1|m+1}^{[NF]} \right)^{-1} = \sum_{i=1}^{N} \left(P_{m+1|m+1}^{[N,i]} \right)^{-1} - (N-1) \left(P_{m+1|m}^{[N]} \right)^{-1}. \quad (30)$$

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 (30).

4. Experiments for fusion of perceptions

The experiments have been carried out with simulated measurement data obtained from virtual reality. The state variables vector $\underline{x}$ and the system dynamics matrix Ψ are given by (31) and (32) respectively. Namely,

$$\underline{x}^T = [x, \dot{x}, y, \dot{y}, \omega]^T \tag{31}$$

where ω is the angular rate and it is estimated during the move. When the robot moves in a straight line, the angular rate becomes zero. In (32) T_s is the sampling time; $w_1, \ldots, w_4$ are due to the linearization of the system dynamics and they represent the derivatives of the system matrix with respect to the angular rate ω. Three salient bending modes are seen in figure 9 with the complete trajectory of the perceptual agent.

$$\Psi = \begin{bmatrix} 1 & \sin\omega T_s/\omega & 0 & (\cos\omega T_s - 1)\omega & w_1 \\ 0 & \cos\omega T_s & 0 & \sin\omega T_s & w_2 \\ 0 & (1-\cos\omega T_s)/\omega & 1 & \sin\omega T_s/\omega & w_3 \\ 0 & \sin\omega T_s & 0 & \cos\omega T_s & w_4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{32}$$

For a straight ahead mode, system matrix becomes

$$\Psi = \begin{bmatrix} 1 & T_s & 0 & 0 & 0 \\ 0 & 1_s & 0 & 0 & 0 \\ 0 & 0 & 1 & T_s & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{33}$$

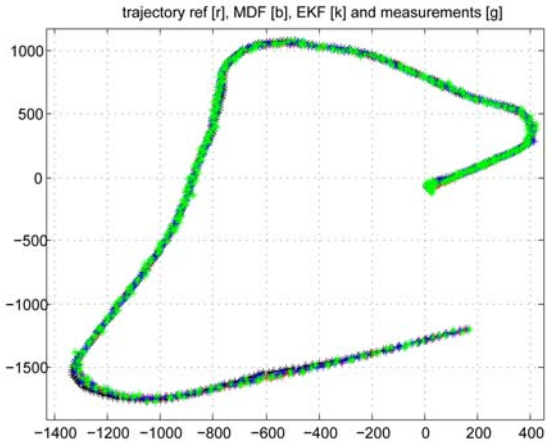

Fig. 9. Robot trajectory, measurement, Kalman filtering and multiresolutional filtering estimation.

In details, there are three lines plotted in figure 10. The solid line is the extended 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. For explicit illustration of the experimental outcomes the same figure with a different zooming range and the zooming power are given in figures 10 and 11 for bending modes and 12 for approximately straight-ahead case.

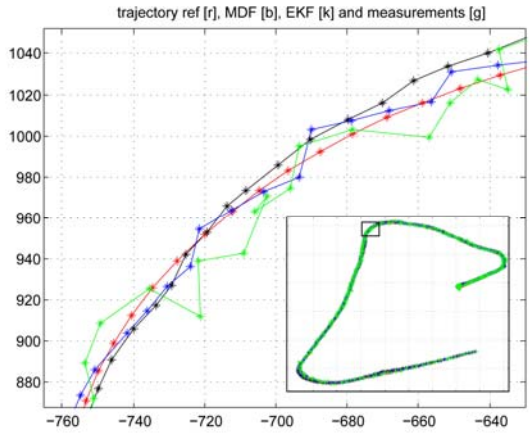

Fig. 10. Enlarged robot trajectory.

In Figs. 10-12, the dark black line is for EKF, smooth line is the reference trajectory (red), light-gray color is for perception measurements green) and dark-gray color is for the multiresolutional dynamic filter (MDF) (blue).

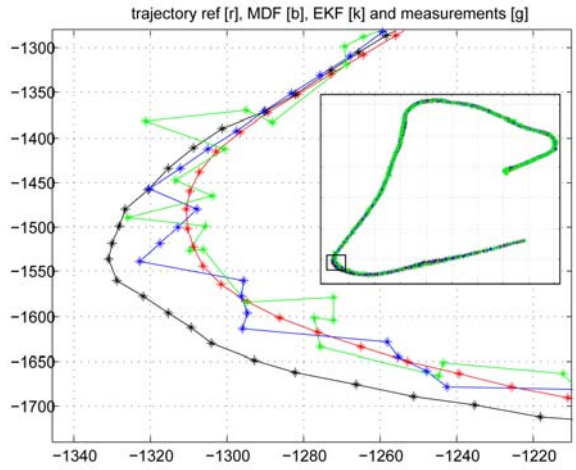

Fig. 11. Enlarged robot trajectory.

From the experiments it is seen that, the Kalman filtering is effective for estimation of the trajectory from perception measurement. Estimation is improved by the multiresolutional filtering. Estimations are relatively more accurate in the straight-ahead mode. It is noteworthy to mention that, the multiresolutional

approach presented here uses *the same* measurements in the lower resolutions. In general case, each sub-resolution can have separate perception measurement from its own dedicated perceptual vision system for more accurate executions. The multiresolutional fusion can still be improved by the use of different data acquisition provisions which play the role of different sensors at each resolution level and to obtain independent information subject to fusion.

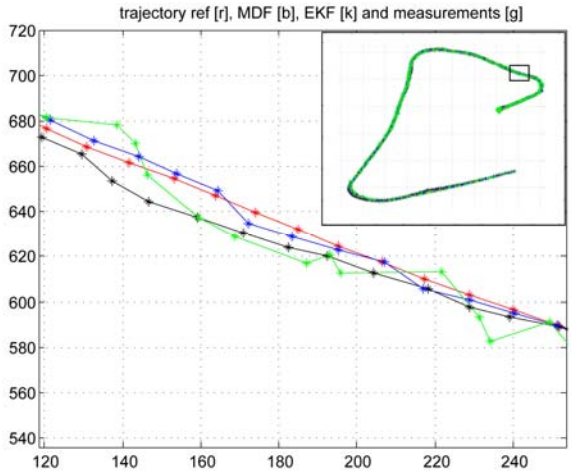

Fig. 12. Enlarged robot trajectory.

5. Discussion and conclusion

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 daily life. However, in professional areas, like architectural design or robotics, its demystification or precise description is necessary for 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 in visual perception. Posit of this research is, that perception is a very complex process including brain processes. In fact, the latter, i.e., the brain processes, about which our knowledge is highly limited, are final, and therefore they are 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 line with our common visual perception experiences, which we exercise every day. The result is used in vision robotics where vision process is accomplished simply by the perception process. Since perception is never quantified before, the application is completely novel in robotics field the particular branch of which is known as perceptual robotics. The novel approach not only performs a vision task in a robot but also it provides it with a simulated human vision. This can be of essential interest in different robotics applications coupled with soft computing technologies. Such applications fall in the category of intelligent robotics. Next to robotics, perception is an important concept subject to consideration in various disciplines as cognitive sciences, Architecture, cybernetics etc. Consequently, the present work in robotics manifests itself as an interdisciplinary research involving several advanced technologies while spanning agent and information technologies.

6. References

[1] Ö. Ciftcioglu, M. S. Bittermann, and I. S. Sariyildiz, "Autonomous robotics by perception," presented at SCIS & ISIS 2006, Joint 3rd Int. Conf. on Soft Computing and Intelligent Systems and 7th Int. Symp. on advanced Intelligent Systems, Tokyo, Japan, 2006.

[2] M. S. Bittermann, I. S. Sariyildiz, and Ö. Ciftcioglu, "Visual Perception in Design and Robotics," *Integrated Computer-Aided Engineering*, vol.14, no.1,pp. 73-91, 2007.

[3] Ö. Ciftcioglu, M. S. Bittermann, and I. S. Sariyildiz, " Further Studies on Visual Perceptionfor Perceptual Robotics," ICINCO 2007 – 4th Int. Conf. on Informatics in Control, Automation and Robotics, 9-12 May, Angers, France, 2007.

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

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

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

[7] R. T. Ogden, *Essential Wavelets for Statistical Applications and Data Analysis*, Birkhauser, Boston, 1997.

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