Intelligent Multi-Sensor Fusion Techniques in Flexible Manufacturing Workcells

Manish Kumar and Devendra P. Garg, *ASME Life Fellow*

Abstract— **This paper advances specific strategies that can be utilized to fuse data from some of the most extensively used sensors in robotic workcells viz. vision sensors and proximity sensors. Vision sensor and proximity sensor are used to obtain the workspace occupancy information. Data from these redundant, yet diverse, sensors have been fused using Bayesian inference to obtain an occupancy grid model of the workspace. In addition, the paper investigates the use of Kalman filtering technique to estimate the external forces acting on robot end-effector utilizing its underlying dynamics and data from force/torque (F/T) sensor mounted on the wrist of the robot. The camera to robot transformation used in the experiment is obtained via a neural network training approach. The proposed strategy to obtain transformation and data fusion is tested and validated in a robotic work cell using one ABB IRB140 six-axis revolute jointed industrial robot fitted with force/torque sensor, proximity sensor and one camera located at the top of the work cell.**

I. Introduction

Modern day robots operate in an inherently uncertain manufacturing world. The uncertainty arises in the perception and modeling of environments, in the manipulation of robot arms and objects, and in the planning and execution of desired tasks. In order for the robots to extend their abilities in uncertain and flexible environments, sensor systems must be developed which can dynamically interpret the observations from the environment in terms of a task to be performed, accounting for this uncertainty and obtaining an accurate model of the robot world. A multi-sensor system must use algorithm or strategy to model the sensor inaccuracies, properly compensate for uncertainties, fuse information from multiple sensory sources and acquire the most accurate model of the environment possible. The algorithms used to fuse data from multiple sensory sources can be classified into three categories: 1) Fusion based on probabilistic methods, 2) Fusion based on least-square techniques, and 3) Fusion based on intelligent methods. The methods employed in this paper are based on Bayesian Technique (probabilistic methods) and Kalman Filtering Technique (least-square techniques). Researches are also underway that make use of intelligent and least-square techniques together to obtain better and more accurate sensor fusion. For example, Sasiadek [1] has proposed one such effort based on adaptive fuzzy logic and extended Kalman filtering.

Vision has played a very important role in modern manufacturing work cells; particularly, in applications such as product assembly and material handling. Vision has also found its application in active feedback control of robots. This application of vision feedback in control loop, generally called visual servoing (see, for example references [2-4]) has been extensively researched. A number of researchers have explored the use of vision information together with force/torque information [5, 6] for grasping and tracking surfaces in robotic work cell.

Manuscript received September 19, 2003. This work was supported by the National Science Foundation under research award number 99-08177.

M. Kumar is a doctoral student in Department of Mechanical Engineering and Materials Science at Duke University, Durham, NC, USA. (email: manish@duke.edu)

D. P. Garg is Professor in Department of Mechanical Engineering and Materials Science at Duke University, Durham, NC, USA. (email: dpgarg@duke.edu)

II. Occupancy Grid

Occupancy grid [7, 8] is a multi-dimensional field (usually of dimension two or three) where each cell (or unit of the grid) stores or represents the probabilistic estimate of the state of spatial occupancy. If the state variable (occupancy, in this case) associated with a cell, C_i, is denoted by $s(C_i)$, then occupancy probability $P[s(C_i)]$ represents the probabilistic estimate of occupation of that particular cell. If $P[s(C_i) = occ] \approx 0$, then the cell is assumed to be empty, while, if $P[s(C_i) = occ] \approx 1$, then the cell is assumed to be occupied.

If a single sensor is used to obtain the occupancy grid, Bayes' Theorem [9, 10] can used in the following manner to determine the state of the cell:

$$P[s(C_i) = occ|z] = \frac{p[z|s(C_i) = occ]P[s(C_i) = occ]}{\sum_{s(C_i)} p[z|s(C_i)]P[s(C_i)]} \quad (1)$$

where z is the sensor measurement. The probability distribution function (p.d.f) $p[z|s(C_i) = occ]$ is dependent on the sensor characteristics and is called the sensor model. The probability $P[s(C_i) = occ]$ is called prior probability mass function and specifies the information made available prior to any observation. This can be extended to incorporate readings from multiple sensors as follows:

$$P[s(C_i)=occ|z_1,z_2]=\frac{p[z_1|s(C_i)=occ]p[z_2|s(C_i)=occ]P[s(C_i)=occ]}{\sum_{s(C_i)}p[z_1|s(C_i)]p[z_2|s(C_i)]P[s(C_i)]} \quad (2)$$

Equation (2) shows measurement from two sensors only that can be extended to any number of sensors. Since the denominator depends only on the measurement (the summation is carried out for all possible values of state: empty and occupied in this case), the probability $P[s(C_i)=occ|z_1,z_2]$ can be expressed as:

$$P[s(C_i)=occ|z_1,z_2]\propto p[z_1|s(C_i)=occ]p[z_2|s(C_i)=occ]P[s(C_i)=occ] \quad (3)$$

which can be optimized by maximizing the right hand side of expression (3).

III. Sensor Models

Sensor modeling is concerned with developing an understanding of the sensed environment [11], the nature of the measurements provided by the sensor, the limitations of the sensor, and probabilistic understanding of the sensor performance in terms of the uncertainties. In addition, sensor model also incorporates optimization of information collection activities and reduction of measurement uncertainties.

The research work presented in this paper makes use of two different sensors (viz. proximity sensor and vision sensor), to obtain the occupancy grid of the work table. The proximity sensor, which is mounted on the wrist of the robot, is of optical variety that has binary (0 or 1) output. However, due to inherent inaccuracies, it continues to detect the object even after the sensor has passed away from the object. This inaccuracy of the sensor was investigated using a small pointed object (screw) and a model of the sensor was obtained. This model (p.d.f, shown in Figure 1) of the sensor provides the probability of occurrence of the object with respect to the distance from the sensor.

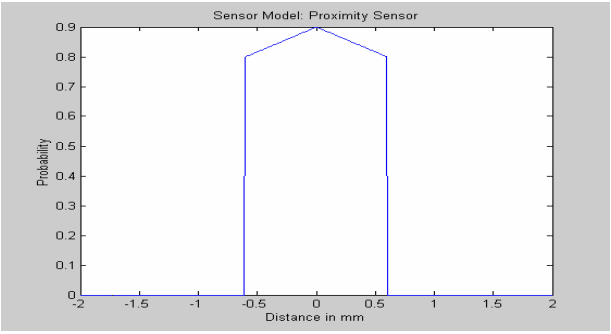

Fig. 1: Proximity Sensor Model

A similar model was obtained for the vision sensor. However, the p.d.f, in this case, was assumed to be Gaussian (see Figure 2). Twenty five data points were obtained which represented the actual spatial coordinates of a point on work table and coordinates obtained from the sensor measurements. These data points were used to estimate the parameter of Gaussian distribution using Maximum Likelihood (ML) method.

Let the posterior distribution be defined by following Gaussian distribution:

$$P(x|z,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{\left\{\frac{-(x-z)^2}{2\sigma^2}\right\}} \quad (4)$$

where x is spatial coordinate to be determined, and z is the spatial coordinate obtained from the measurement. The parameter σ, the standard deviation, can be determined using the ML method in the following way.

$$P(\sigma|x,z)\propto P(x|z,\sigma)$$

Hence,

$$P(\sigma|x,z)=\arg\max P(x|z,\sigma) \quad (5)$$

If there are 'n' data points, then equation (5) can be extended to:

$$P(\sigma|x,z)=\arg\max[P(x_1|z_1,\sigma)P(x_2|z_2,\sigma).....P(x_n|z_n,\sigma)] \quad (6)$$

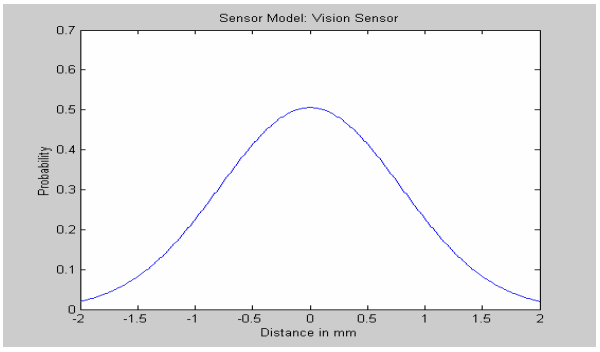

Fig. 2: Vision Sensor Model

IV. Sensor Fusion

The robot, with proximity sensor mounted on its wrist, scans over the work table, storing the proximity data in a two dimensional matrix. This matrix defines the occupancy grid obtained from proximity sensor alone. Similarly, the image obtained from the camera located above the work cell is processed to obtain the location of objects which is then mapped into similar occupancy grid (this grid is represented in robot world frame). In this paper, the mapping from image coordinates to robot world coordinates has been achieved with the help of neural network [12, 13]. Inputs to the neural networks are two-dimensional image coordinates and outputs are the two-dimensional coordinates in robot frame. Twenty five (25) data points were fed to a three-layered neural network having one input layer, one output layer and one hidden layer. Input and output layers had two nodes each and hidden layer had ten nodes. The number of layers and

number of nodes in hidden layer were optimized using trial and error. All of the nodes had linear activation function. The mean square error for the trained network was found to be 1.6327 mm^2. The information stored in these two occupancy grids needed to be fused in order to provide a better probabilistic estimate of the occurrence of the object on the work table.

In the previous section, the model of the sensor was obtained that provided the probability of occurrence of object as a function of distance from measurement, given that the object was detected by the sensor. This model of sensor can be used to fuse the information from the neighboring cells. In a continuous sense, this can be represented by the following integral:

$$P(s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} zP(x,y|z)dxdy \quad (7)$$

where *P(s)* is the probability of occurrence at *x=0* and *y=0*, and *z* is the sensor measurement. In the occupancy grid formulation, the above integral can be replaced by a summation operation carried over all the cells.

$$P(s) = \sum_i zP(s(C_i)) \quad (8)$$

For the occupancy grid developed in this research, each cell is of size 10mm X 10mm. From the sensor model, it follows that only the neighboring cells affect the probability distribution of any particular cell. The fused probability estimate is the average of the values obtained for proximity and vision sensor.

V. END EFFECTOR FORCE/MOMENT ESTIMATION

The dynamics of the process can be modeled as follows. Figure 3 shows a schematic diagram of a gripper, and the forces and torques acting on it. Let $\vec{c}$ be the position vector of the Center of Gravity (C.G.) of the gripper with respect to the world reference frame, $\vec{f}_s$ and $\vec{n}_s$ be the force and moment sensed by the force/torque sensor, $\vec{f}_e$ and $\vec{n}_e$ be the force and moment experienced at the tip of the end effector, $\vec{r}_s$ be the vector from C.G. of the gripper to the force/torque sensor, and $\vec{r}_e$ be the vector from C.G. of the gripper to the center of the end effector

In addition, let m_G be the mass of the gripper, I_G be the moment of inertia of the gripper, ω be the rotational velocity of the gripper, and, θ be the angular position of the gripper. Then, dynamic force and moment equations are:

$$m_G \ddot{\vec{c}} = \vec{f}_s + \vec{f}_e + m_G \vec{g} \quad (9)$$

$$I_G \dot{\vec{\omega}} + \vec{\omega} \times I_G \vec{\omega} = \vec{n}_s + \vec{r}_s \times \vec{f}_s + \vec{n}_e + \vec{r}_e \times \vec{f}_e \quad (10)$$

These relations have been obtained with the use of Newton-Euler formulations and the effects of Coriolis forces have been ignored.

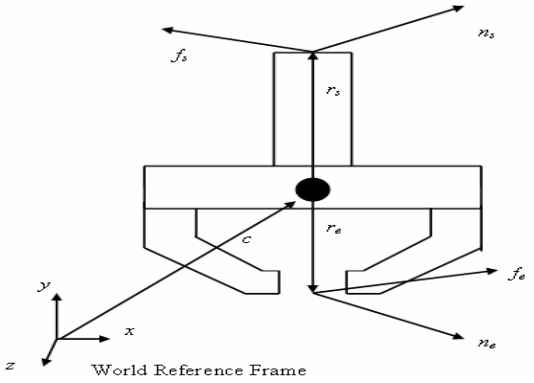

Fig. 3: Forces and Moments Acting on Gripper

A. Kalman Filtering

Kalman filter has traditionally been used extensively in the solution of tracking, estimation and signal extraction problems. In robotics, Kalman filter has been used to estimate contact forces, uncertainties in robotic contacts [14, 15], and external forces acting on the gripper [16, 17]. This section provides a brief description about Kalman filter and its use in robotic workcell.

Kalman filtering [18-20] is an optimal recursive data processing algorithm that is based upon state space concepts. The recursive nature of the algorithm makes it suitable for systems without large data storage capacities. The variables estimated using Kalman filter can be shown to be statistically optimal because the approach uses: i) knowledge of system and measurement device dynamics, ii) the statistical description of system noise, measurement errors, and uncertainty in dynamic models, and iii) all available information about the initial conditions of variables of interest. If the system can be described via a linear model, and both the system and sensor error can be modeled as white Gaussian noise, then Kalman filter provides unique statistically optimal estimates for the fused data.

Let $x \in R^n$ be the state of discrete time controlled process that needs to be estimated. Let the process be governed by the linear stochastic difference equation:

$$x_k = Ax_{k-1} + Bu_k + w_{k-1} \quad (11)$$

with a measurement $z \in R^m$, i.e.,

$$z_k = Hx_k + v_k \quad (12)$$

where w_k and v_k are the random variables that represent the process and measurement noise respectively. These noises are assumed to be Gaussian, white and independent of each other. *A* is (n x n) matrix which relates the state at previous step *(k-1)* to the state at current step *(k)*, $u \in R^l$ is the control input and '*B*' is (n x l) matrix relating this

control input to current state. The (m x n) matrix H in the measurement equation (12) relates the state to the measurement z_k.

$$p(w) \approx N(0, Q) \quad (13)$$

$$p(v) \approx N(0, R) \quad (14)$$

where Q and R represent the process and measurement noise covariance matrices respectively.

Let $\hat{x}_k^-$ be a priori state estimate at step k and $\hat{x}_k$ be *a posteriori* state estimate. The *a priori* and *a posteriori* estimate errors can be defined as:

$$e_k^- \equiv x_k - \hat{x}_k^- \quad (15)$$

$$e_k = x_k - \hat{x}_k \quad (16)$$

The current state is estimated using the following set of equations:

Time update equations:

$$\hat{x}_k^- = A\hat{x}_k + Bu_k \quad (17)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (18)$$

Measurement update equations:

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (19)$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \quad (20)$$

$$P_k = (I - K_k H)P_k^- \quad (21)$$

where P_k and P_k^- are *a posteriori* and *a priori* estimate error covariance matrices.

B. End Effector Force Estimation

The process equation is:

$$X(n) = AX(n-1) + BU(n) + w(n) \quad (22)$$

and the measurement equation is

$$Y(n) = HX(n) + v(n) \quad (23)$$

where *n* denotes the time instant, and $X(n)$ is the state vector representing the force to be measured (f_e). From the dynamics of the system, the following values have been derived:

$A = [1]$, $B = [0.005]$ and $H = [-1]$

The noise covariance matrices are given by:

$$Q = E\{w(n)w(n)^T\} = [0.001], \text{ and} \quad (24)$$

$$R = E\{v(n)v(n)^T\} = [0.001]. \quad (25)$$

Noise covariance matrices Q and R represent a measure of confidence in the process equation (dynamics) and the measurement respectively. These matrices have been chosen based on experimentation, and equal value of both of these matrices represents equal confidence in dynamics of the process and the measurement. $U(n)$ is the vector representing the acceleration (mm/sec^2) of center of gravity of the gripper in three directions, and $Y(n)$ is the force measured by the sensor in three directions. For the experiment testbed (Figure 4) available in the Robotics and Manufacturing Automation (RAMA) Laboratory, the mass of the gripper is 5 kg and the matrix B has been chosen based on the mass. All the vectors (force, acceleration) have been transformed to robot coordinate frame.

C. End Effector Moment Estimation

The process and measurement equations remain the same as the force estimation equations:

$$X(n) = AX(n-1) + BU(n) + w(n) \quad (26)$$

$$Y(n) = HX(n) + v(n) \quad (27)$$

The various matrices, derived from dynamics of the process, are:

$A = [1]$, $B = [1]$ and $H = [-1]$

The noise covariance matrices are given by:

$$Q = E\{w(n)w(n)^T\} = [0.001] \text{ and} \quad (28)$$

$$R = E\{v(n)v(n)^T\} = [0.001] \quad (29)$$

$U(n)$, in equation (26), is an input vector derived from equation (10) and is given by following equation:

$$U(n) = I_G \dot{\vec{\omega}} + \vec{\omega} \times I_G \vec{\omega} - \vec{r}_s \times \vec{f}_s - \vec{r}_e \times \vec{f}_e \quad (30)$$

The dynamics of the process reveals that relationship between the quantity to be estimated and measured quantity is linear. Added to that, the difference equations (22) and (26) are assumed to represent a linear signal process. Hence, Kalman filter, rather than extended Kalman filter, is used to estimate the forces and moments acting on the tip of the end effector.

Fig. 4: Experimental Setup showing Robots and Objects in the Workcell

VI. Results and Discussions

A. Occupancy Grid

The robot, with proximity sensor attached to its wrist, scanned over the work table and proximity sensor reading were recorded. A three-dimensional plot showing the

proximity sensor measurements is shown in Figure 5. The x and y axes represent the actual coordinates in robot world frame, while the z-axis represents the sensor reading.

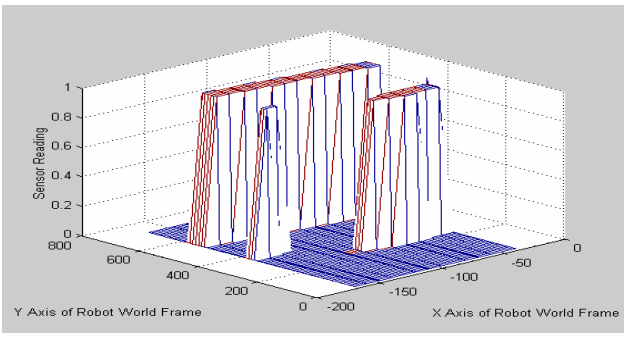

Fig. 5: Proximity Sensor Measurement Plotted against Robot World Coordinates

An image of the work space (Figure 6) was taken using the vision sensor and it was processed to reveal the location of the objects. With the help of appropriate transformations, the locations of the objects were transformed to robot world coordinates. Figure 7 shows another three-dimensional plot (similar to Figure 5 obtained via proximity sensor) that was obtained from the vision data.

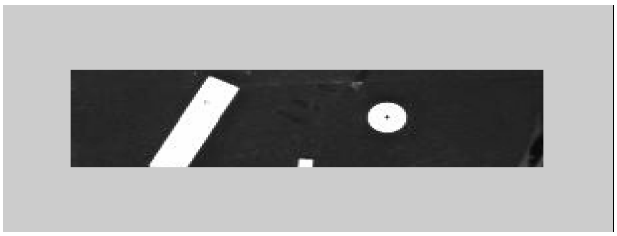

Fig. 6: Image of the Work Table from the Vision Sensor Located above the Work Cell

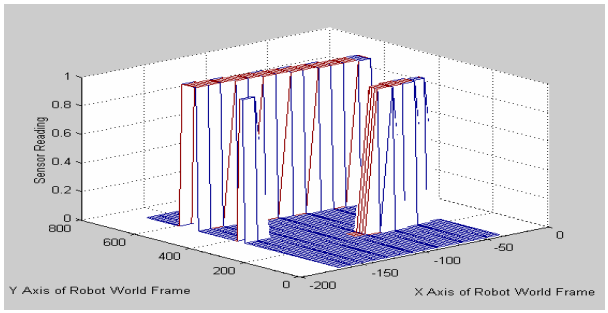

Fig. 7: Vision Sensor Measurement plotted against Robot World Coordinates

The strategies described in previous sections were used to fuse data from the vision and the proximity sensors to obtain the fused occupancy grid. Figure 8 shows the three-dimensional plot of the fused grid. Since the aspect ratio of x and y axes are not the same in figures 5 and 7, the geometric features of the objects (such as roundness, squareness etc.) are not revealed. Figure 8 is a three dimensional plot showing the different objects when the aspect ratio of x and y axes are same. This figure reveals the geometric feature of the objects more accurately. The occupancy grid obtained from fusion of the vision and proximity data compensates for the uncertainties inherent in individual sensors and provides more accurate probabilistic information about the occurrence of object on the work table. The size of the grid for the above simulation has been chosen to be 10 mm X 10 mm, which can be decreased to obtain a more accurate occupancy profile. The sensor models (Figures 1 and 2) show that the probability distribution is limited to small region (1 mm in case of proximity sensor and approximately 5 mm in case of vision sensor). Hence, grid size of that magnitude would yield accurate results. However, making grid size smaller would adversely affect the computing time required. Hence, grid size should be chosen based on sensor capabilities, desired accuracy and computing time requirements.

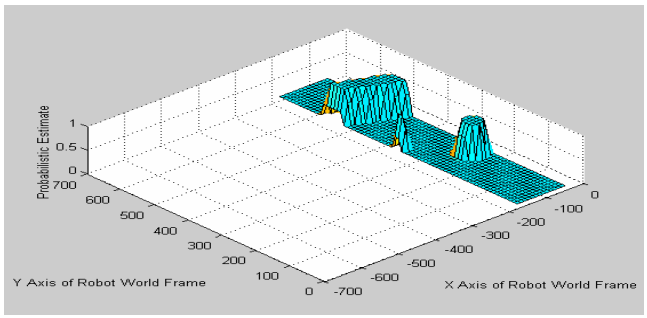

Fig. 8: Three-Dimensional Representation of the Occupancy Grid with Equal Aspect Ratio

B. Kalman Filtering

Experiment was carried out to study the performance of Kalman filter in estimating the forces and moments acting on the end effector. The experiment consisted of repeatedly moving the robot linearly between two points. The robot, in this experiment, was given command to move back and forth a distance of 700 mm along Y-axis. The maximum velocity attained in the movement was constrained to be 150 mm/sec. The position data of the robot end effector were obtained from the ABB controller, which were differentiated to obtain velocity and acceleration of the end effector. Forces were applied externally in several directions at different points of the gripper during the experiment by an operator.

Figure 9 shows the forces measured by force/torque sensor as compared to forces estimated by Kalman filtering technique described above. As expected, the forces measured by the sensor are in opposite direction with

respect to the applied force. Similar representation was obtained comparing the measured and sensed moments. The forces and moments estimated by the filter were in line with the direction and magnitude of forces and moments applied externally.

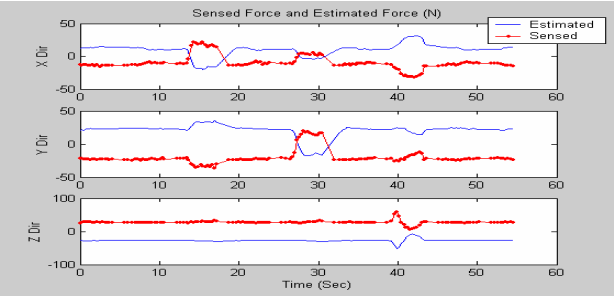

Fig. 9. Forces Sensed and Estimated in Three Directions

C. Occupancy Grid and Force Sensing

Sensory information, occupancy grids and force/torque data can prove to be highly complementary. For example, occupancy grids can provide approximate information about position, orientation and geometric profile of objects on the work table. This knowledge can be fused with the force/torque information when robot makes contact with the object. The direction and magnitude of forces/torques can reveal important characteristics about the surface profile. This collective cognition can be used to accomplish several goals, such as surface tracking, and grasping irregular objects, that would be very difficult to achieve otherwise.

VII. Conclusion

This paper presents strategies to convert practical data obtained from some of the commonly used sensors in robotic systems into a form that can be used directly to control and coordinate robots. The first part of the paper is dedicated to obtaining occupancy grid of the work table via the use of proximity and vision sensor. The strategy formulated has been tested with the help of real world sensors and results have been presented. The second part of the paper presents a technique based on Kalman filtering to extract the forces and moments acting on the tip of the manipulator based on the measurements from the force/torque sensors. Since the force/torque sensors are usually mounted on the wrist of the robot, and not at the location where the measurement is desired, dynamics of the gripper comes into play. The paper discusses this dynamics and uses it to extract the values of external forces and moments.

References

[1] Sasiadek, J.Z., "Sensor Fusion", *Annual Reviews in Control*, Vol. 26, 2002, pp. 203-228.

[2] Han, S. H., See, W. H., Lee, J., Lee, M. H., and Hashimoto, H., "Image-Based Visual Servoing Control of a SCARA Type Dual-Arm Robot", *Proceedings of the 2000 IEEE International Symposium on Industrial Electronics*, Vol. 2, 4-8 Dec. 2000, pp. 517 –522.

[3] Kragic, D., and Christensen, H. I., "Integration of Visual Cues for Active Tracking of an End-Effector", *Proceedings of International Conference on Intelligent Robots and Systems*, Vol. 1, 17-21 Oct. 1999, pp. 362 –368.

[4] Asada, M., Tanaka, T., and Hosoda, K., "Visual Tracking of Unknown Moving Object by Adaptive Binocular Visual Servoing", *Proceedings of IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems*, 15-18 Aug. 1999, pp. 249 –254.

[5] Allen, P. K., Miller, A.T., Oh, P. Y., and Leibowitz, B. S., "Integration of Vision and Force Sensors for Grasping", *Proceedings of IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems*, 1996, pp. 349-356.

[6] Baeten, J., Bruyninckx, H., and De Schutter, J., "Combining Eye-in-Hand Visual Servoing and Force Control in Robotic Tasks Using the Task Frame", *Proceedings of IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems*, 15-18 Aug. 1999, pp.141 –146.

[7] Abidi M. and Gonzalez, R., "Data Fusion in Robotics and Machine Intelligence", *Academic Press, Inc.*, 1992.

[8] Puente, E.A., Moreno, L., Salichs, M.A., and Gachet, D., "Analysis of Data Fusion Methods in Certainty Grids Application to Collision Danger Monitoring", *Proceedings of IEEE International Conference on Industrial Electronics, Control and Instrumentation*, 28 Oct. – 1 Nov, 1991, pp 1133 – 1137.

[9] Press, S.J., *Bayesian Statistics: Principles, Models and Applications*, John Wiley and Sons, 1989.

[10] Luo, R. and Su, K., "A Review of High-Level Multisensor Fusion: Approaches and Applications", *Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems,* 1999, pp. 25-31.

[11] Manyika, J. and Durrant-Whyte, H., *Data Fusion and Sensor Management: a Decentralized Information-Theoretic Approach*, Ellis Howard Limited, 1994.

[12] Garg, D. P., and Kumar, M., "Camera Calibration and Sensor Fusion in an Automated Flexible Manufacturing Multi- Robot Work Cell", *Proceedings of the American Control Conference*, Vol. 6, 2002, pp. 4934 –4939.

[13] Nagachaudhuri, A., Thint, M. and Garg, D. P., "Camera Robot Transform for Vision Guided Tracking in a Manufacturing Work Cell", *Journal of Intelligent and Robotic Systems,* Vol. 5, 1992, pp. 283-298.

[14] Katupitiya, J., Dutre, S., Demey, S., De Geeter, J., Bruyninckx, H., and de Schutter, J., "Estimating Contact and Grasping Uncertainties Using Kalman Filters in Force Controlled Assembly", *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 2, Nov 1996, pp. 696 – 703.

[15] Chua, A., and Katupitiya, J., "Kalman Filters for the Identification of Uncertainties in Robotic Contact", *Proceedings of the 37th IEEE Conference on Decision and Control*, Vol. 2, Dec 1998, pp. 2013 – 2018.

[16] Uchiyama, M., Yokota, M., and Hakomori, K., "Kalman Filtering 6-Axis Robot Wrist Force Sensor Signal", *Proceedings of International Conference on Advanced Robotics*, Tokyo, Japan, 1985, pp. 153-160.

[17] Uchiyama, M., and Kitagaki, K., "Dynamic Force Sensing for High Speed Robot Manipulation Using Kalman Filtering Techniques", *Proceedings of 28th Conference on Decision and Control*, Tampa, Florida, December 1989, pp. 2147 – 2152.

[18] Maybeck, P. S., *Stochastic Models, Estimation and Control*, Volume 1, Academic Press Inc., 1979.

[19] Hayes, M. H., *Statistical Digital Signal Processing and Modeling*, John Wiley & Sons Inc., 1996

[20] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems", *Transactions of the ASME, Journal of Basic Engineering*, Vol. 82, Ser. D, 1960, pp. 35 – 45.