Object Boundary Detection Using An Optimal Data Association Scheme

Jung-Gu Kim and Hong Jeong

Abstract

In target tracking area, the data association plays an important role and has been studied extensively. In this paper, after defining the data association as a constrained optimization, we introduce a new energy function and thereby an efficient realization of neural networks.

As an application, this algorithm is used to detect object boundaries in IR images. The problem is that the IR image noisy, the shape of the object is variable, and the positions of the end points are not predictable. The performance of this algorithm is discussed with the experimental results.

I. Introduction

The primary purpose of a multi-target tracking system [1], [3], [7], [8], [9], [10] is to track as accurately as possible the moving target, guided by some data measured in the field of view. Naturally, the performance of this system is prone to errors such as the clutter, false detection, missed target, and uncertain target location.

Recently, Sengupta and Iltis [8] successfully used the constrained minimization technique to compute posterior probabilities in JPDA(Joint Probabilities Data Association). According to this scheme, the posterior β_j^t (for $j \neq 0$) is the probability that measurement j originates from target t. Similarly, β_0^t is the probability that none of the received measurements originates from target t. They proposed Hopfield network to approximately compute the β_j^t and called this scheme NNPDA(Neural Network Probabilistic Data Association). In particular, β_j^t is approximated by the output voltage X_j^t of a neuron in an $N \times (T+1)$ array of neurons, where N and T are respectively the number of measurements and the number of targets.

However, the energy function in [8] is inefficient [7] in that it does not adopt the full advantages of the natural constraints. Since the value of β_j^t in the original JPDA is inconsistent with X_j^t of [8], these dual assumptions of no two returns from the same target and no single return from two targets should be used only in the generation of the feasible data association hypotheses, as pointed out in [1]. This resulted from misinterpretation of the properties of JPDA which the network was supposed to emulate.

We propose a new neural network scheme which reflects better the natural constraints of the multi-target tracking problem [6]. As an application to computer vision, we applied the algorithm to the detection of object boundaries in IR image. Especially each image contains one thick plate that is obtained in front of the steel mill. The shape of the steel plates are variable depending on the rolling conditions.

The organization of the rest of this paper is as follows. In § II, we define the data association as a constrained optimization and as one of optimal solutions we suggest a MAP estimate. The various energy terms in the posterior that is used in the MAP are derived in § III. A neural network is derived in § IV, that efficiently converges to a minimizer of the energy function. In § V, we apply this algorithm to detect object boundary in infrared image. Finally, in § VI we showed experimental results.

II. A MAP Estimate for Optimal Data Association

Suppose that there are T targets and N number of measurement data in some appropriate gates at some time instance. The problem is that when a set of measurement data is received, it has to be determined that which

Manuscript received October 4, 1995; accepted December 12, 1995.
The authors are with Department of Electrical Engineering, Pohang University of Science and Technology, Pohang 790-784, Korea

measurement is to be associated with which target. The relationships of these measurements and targets are the problem of multiple target tracking which is best described by $N\times(T+1)$ *validation matrix*[3]

$$\Omega=[\omega_{ij}]=\begin{bmatrix}\omega_{10} & \omega_{11} & \omega_{12} & \cdots & \omega_{1T}\\ \omega_{20} & \omega_{21} & \omega_{22} & \cdots & \omega_{2T}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ \omega_{N0} & \omega_{N1} & \omega_{N2} & \cdots & \omega_{NT}\end{bmatrix}. \quad (1)$$

The elements of the validation matrix are bound between 0 and 1: $\omega_{ij}=1$ if the ith measurement originated from the jth target or object boundary, and $\omega_{ij}=0$ otherwise. And the first column of this matrix is considered as false detection and others as targets. All rows are regarded as measurement data.

An example of typical measurement data distribution within two different gates of two targets is shown in Fig. 1.

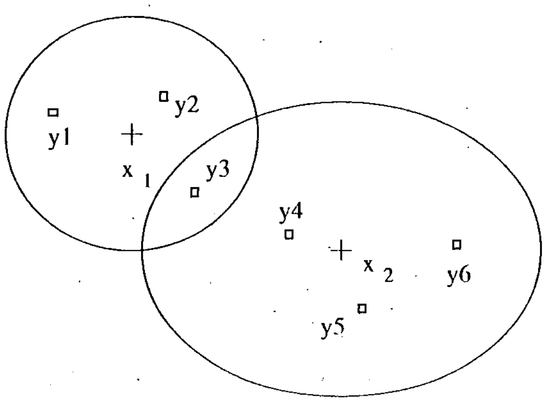


Fig. 1. The distribution of measurement data within two overlapping gates.

In this figure, x_i and y_j represent the center of the validation gate[3] for the ith target and the jth measurement, respectively. Measurements y_1, y_2 and y_3 are data for the first target and y_3, y_4, y_5 and y_6 for the second target. Generally, a set of gates can be overlapped if there is at least one common measurement in their scopes.

For the purpose of solving the problem of a large number of measurements and targets, one must find the *data association matrix*. The data association matrix for different gates and measurements can be formulated based on the validation matrix (1). In the case of Fig. 1, one of feasible candidate data association matrices is

$$\omega=\begin{bmatrix}1&0&0\\0&1&0\\1&0&0\\1&0&0\\1&0&0\\0&0&1\end{bmatrix} \quad (2)$$

This means that the measurement y_2 must belong to the first target, y_6 to the second target and others to the clutter.

The generation of the association matrix leads to a combinatorial problem where the number of data association hypothesis increases exponentially with the number of target and the number of measurement. We consider that the task of finding this feasible association matrix is a constrained optimization problem. And we define this problem by the MAP estimation formula:

$$\begin{cases}\text{Given} & z=(y,x,\hat{\omega}),\\ \text{find} & \omega^*,\\ \text{suchthat} & \omega^*=\arg\max_{\omega}P(\hat{\omega},\omega|z),\end{cases} \quad (3)$$

where y, x, and $\hat{\omega}$ respectively represent measurement data in a validation gate, center of the gate, and initial validation matrix. And $P(\cdot)$ denotes a posterior pdf and ω^* is an optimal estimator of ω for the given z.

III. Determining Energy Functions for the MAP

To solve (3), we need an exact representation of the posterior probability. According to Bayes rule, the posterior becomes

$$P(\omega|y,x,\hat{\omega})=\frac{P(\hat{\omega}|\omega)P(y,x|\omega)P(\omega)}{P(y,x,\hat{\omega})}. \quad (4)$$

All the pdfs in (4) are assumed to have the Gibbs distribution[2]:

$$\begin{cases}P(\omega|y,x,\hat{\omega}) &= \frac{1}{Z_1}e^{-\frac{1}{k\tau}E(\omega|y,x,\hat{\omega})},\\ P(\hat{\omega}|\omega) &= \frac{1}{Z_2}e^{-\frac{1}{k\tau}E(\hat{\omega}|\omega)},\\ P(y,x|\omega) &= \frac{1}{Z_3}e^{-\frac{1}{k\tau}E(y,x|\omega)},\\ P(\omega) &= \frac{1}{Z_4}e^{-\frac{1}{k\tau}E(\backslash omega)}.\end{cases} \quad (5)$$

Here Z_1, Z_2, Z_3 and Z_4 are all partition functions. And k, τ and $E(\cdot)$ denote a temperature constant, Boltzmann constant in thermal dynamics for controlling the convergence rate and energy function, respectively.

We assume the following properties for the energy function:

$$E(\omega|y,x,\hat{\omega})=E(\hat{\omega}|\omega)+E(y,x|\omega)+E(\omega). \quad (6)$$

The first term $E(\hat{\omega}|\omega)$ in (6) denotes the matching relationship between initial association matrix and optimal association matrix, and their distance is usually modeled by a Gaussian distribution. The second term $E(y,x|\omega)$ stands for the relation between measurement data and the center of the gate of object boundary at some time instance with the given optimal data association matrix. And the last term $E(\omega)$ has the meaning of prior knowledge of the association matrix itself.

The energy for the first term therefore can be represented by

$$E(\hat{\omega}|\omega)=\alpha\sum^{N_{i=1}}\sum^{T_{j=1}}(\hat{\omega}_{ij}-\omega_{ij})^2, \qquad (7)$$

where α is a constant.

Now let's consider an arbitrary trajectory with data y_i within validation gate as shown in Fig. 2.

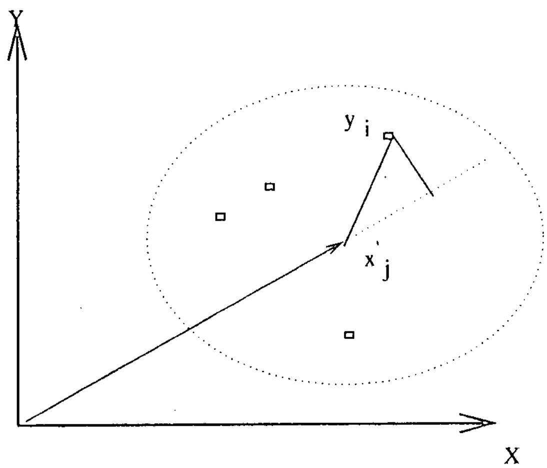


Fig. 2. Distance between measurement data and trajectory.

In this figure x_j denotes the center of the gate of target trajectory or object boundary estimated at previous time step. Using simple vector operation, the direction of arbitrary trajectory and perpendicular distance between data and the trajectory can be calculated. Here, we assume that the energy for measurement located close to the trajectory has to be low than that of data located far from the trajectory.

Therefore the energy function for the second term can be defined as

$$E(y,x|\omega)=\beta\sum_{i=1}^{N}\sum_{j=1}^{T}\frac{(x_i y_j-y_i x_j)^2}{x^{2_i}+y^{2_i}}\omega_{ij}, \qquad (8)$$

where x_i and y_i are respectively the x and y components of the measurement, x_j and y_j are respectively the x and y components of the center of validation gate, and β is a constant.

Only one measurement data must be associated with a specific target at each time. Therefore the energy must be low for only those solutions that produce a single 1 in each column and a single 1 in rows from the second column in association matrix. That is,

$$\begin{cases}\sum_{i=1}^{N}\omega_{ij}=1 & \forall j\neq 0,\\ \sum_{j=0}^{T}\omega_{ij}=1 & \forall i\end{cases}. \qquad (9)$$

The following energy function satisfies these constraints:

$$E(\omega)=\lambda\sum_{i=1}^{N}(\sum_{j=0}^{T}\omega_{ij}-1)^2+\gamma\sum_{j=1}^{T}(\sum_{i=1}^{N}\omega_{ij}-1)^2. \qquad (10)$$

Here, the parameters λ and γ are constants.

Substituting (7), (8), and (10) into (6) yields

$$\begin{aligned}E(\omega|y,x,\hat{\omega})=&\ \alpha\sum_{i=1}^{N}\sum_{j=1}^{T}(\hat{\omega}_{ij}-\omega_{ij})^2\\ &+\beta\sum_{i=1}^{N}\sum_{j=1}^{T} frac(x_i y_j-y_i x_j)^2\ x^{2_i}+y^{2_i}\,\text{omega}_{ij}\\ &+\lambda\sum_{i=1}^{N}(\sum^{T_{j=0}}\omega_{ij}-1)^2+\gamma\sum_{j=1}^{T}(\sum_{i=1}^{N}\omega_{ij}-1)^2.\end{aligned} \qquad (11)$$

IV. Minimizing the Energy by Neural Network

Given the total energy (11), we will derive an efficient solution that uses Hopfield network[5]. A Hopfield network which has $N\times(T+1)$ neurons has the following energy function.

$$E=-\frac{1}{2}\sum_i\sum_k\sum_j\sum_l w_{ikjl}\,\omega_{ij}\,\omega_{kl}-\sum_i\sum_j I_{ij}\,\omega_{ij}, \qquad (12)$$

where w_{ikjl} is the connection weight between ikth neuron and jlth neuron and I_{ij} is the external input to the ijth neuron. To find the connection weight and the external input, we rearranged the energy function as

$$\begin{aligned}E=&\sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{j=0}^{T}\sum_{l=0}^{T}\{\alpha\delta_{ik}\delta_{jl}(1-\delta_{0j})+\lambda\delta_{ik}+\gamma\delta_{jl}(1-\delta_{0j})\}\omega_{ij}\omega_{kl}\\ &-\sum_{i=1}^{N}\sum_{j=0}^{T}\{2\alpha(1-\delta_{0j})\,\hat{\omega}_{ij}-\beta(1-\delta_{0j})\frac{(x_i y_j-x_j y_i)^2}{(x_j^2+y_j^2)}+2\lambda+2\gamma(1-\delta_{0j})\}\omega_{ij}\\ &+\sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{j=0}^{T}\sum_{l=0}^{T}\alpha\delta_{ik}\delta_{jl}(1-\delta_{0j})\,\hat{\omega}_{ij}\,\hat{\omega}_{kl}+\lambda N+\gamma T\end{aligned} \qquad (13)$$

where $\sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{j=0}^{T}\sum_{l=0}^{T}\alpha\delta_{ik}\delta_{jl}(1-\delta_{0j})\,\hat{\omega}_{ij}\,\hat{\omega}_{kl}$ and $\lambda N+\gamma T$ are constant.

Comparing (12) with (13), we can obtain the value of the weights and inputs:

$$\begin{cases}w_{ikjl}=-2\{\alpha\delta_{ik}\delta_{jl}(1-\delta_{0j})+\lambda\delta_{ik}+\gamma\delta_{jl}(1-\delta_{0j})\},\\ I_{ij}=\{2\alpha(1-\delta_{0j})\,\hat{\omega}_{ij}-\beta(1-\delta_{0j})\frac{(x_i y_j-x_j y_i)^2}{(x_j^2+y_j^2)}+2\lambda+2\gamma(1-\delta_{0j})\}\end{cases} \qquad (14)$$

Based on (14), we can draw the overall architecture of Hopfield network as in Fig. 3. In this figure, the inputs for the first column of neurons are

$$I_{i0}=2\lambda, \qquad (15)$$

and inputs for others are

$$I_{ij\neq 0}=2\alpha\,\hat{\omega}_{ij}-\beta\frac{(x_{iy_i}-x_{jy_i})^2}{x_j^2+y_j^2}+2(\lambda+\gamma). \qquad (16)$$

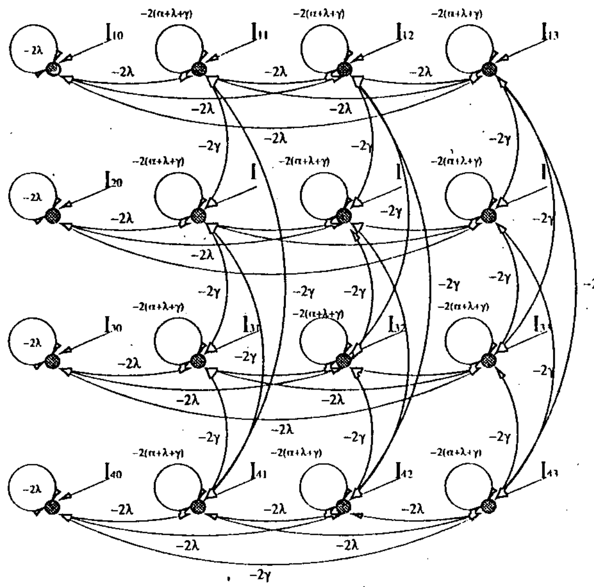


Fig. 3. Hopfield neural network architecture for data association.

Notice that there is no connection in the vertical direction of the first column. However, the neurons on the same rows and columns are connected so that they may inhibit each other. All the neurons receive inputs that contain observation data.

V. Application to Boundary Detection

So far, we have derived a neural network that solves the data association problem. In order to apply this algorithm to boundary detection, we consider the boundary as target and the edges as measurement data. Then, the overall boundary detection system becomes Fig.4 that consists of *acquisition* and *tracking* system.

The operation is as follows. At first, the acquisition system provides some starting points. Then, the tracking system searches for the next point of the object boundary. To accomplish this task, the tracking system first predicts, usually by Kalman filter, the region that may contain the candidates of the next boundary points. Within this region, called *gate*, boundary points must be obtained by some operators of edge detection. After then, the data association system chooses one of the measurement data as a boundary point. The loop repeats until some termination condition is reached.

This process was shown in Fig.5.

Starting from the initial site (denoted by a filled square), the Kalman filter predicts a gate whose center is denoted by a cross.

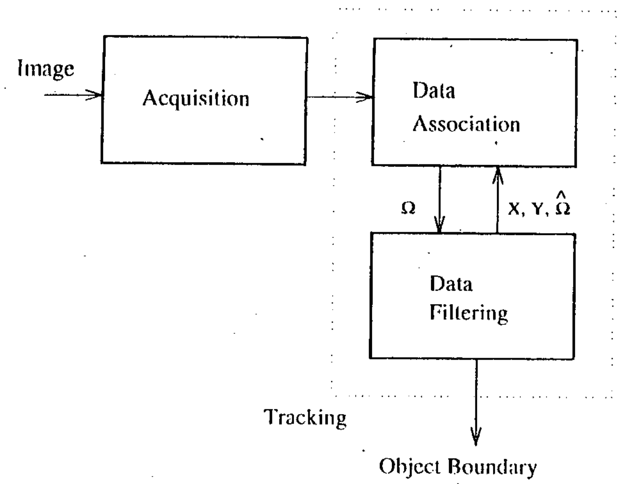


Fig. 4. A boundary detection system.

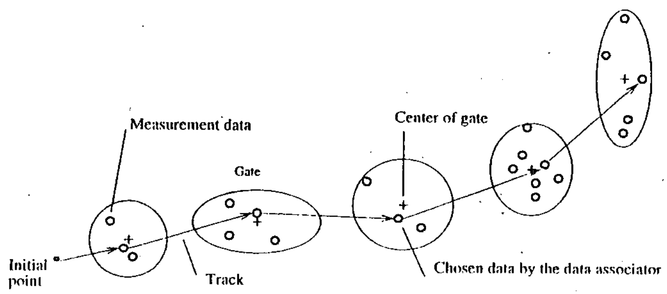


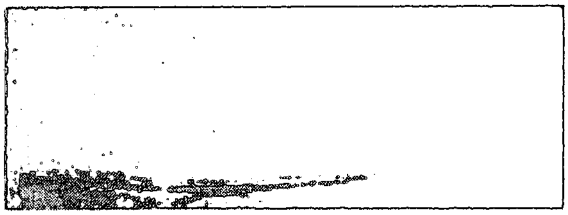
Fig. 5. An example for boundary detection.

Within the range of the gate, we obtain some measurement data of the image that possibly denote boundary points. In this paper, we used a Sobel edge operator [4] for providing the measurement data. Among the many measurement data, the purpose of the data associator is to choose only one as a boundary point. Starting from the newly chosen point, the Kalman filter predicts another gate. As a result of the iterative process, a track represented by a solid line is obtained.

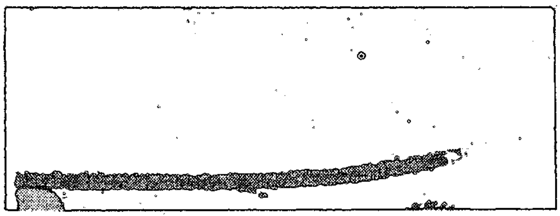
VI. Experimental Results

To test our algorithm, we applied it to two test IR images that contain hot plates. The infra-red image is like an ordinary gray scale image other than that it is relatively free from background noise. The goal of this experiment is to detect the upper and lower boundaries of the plate that can be used later for calculating curvature of the plate by means of some curve-fitting method.

The 210 × 550 test images which have been taken from mono *CCD* camera and IR bandpass filter with 1050 nm center frequency are shown in Fig.6. Each image contains one hot plate that is over 750 °F.



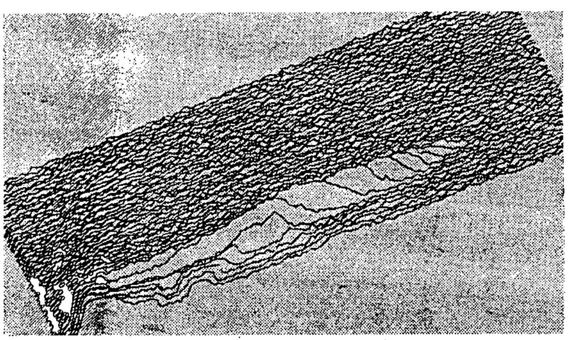
(a)



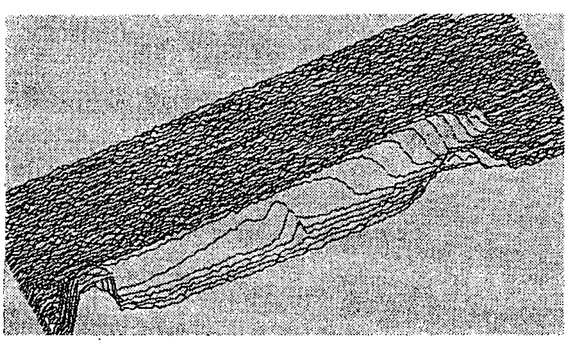
(b)

Fig. 6. (a) and (b) 210 × 550 two test images of hot plate.

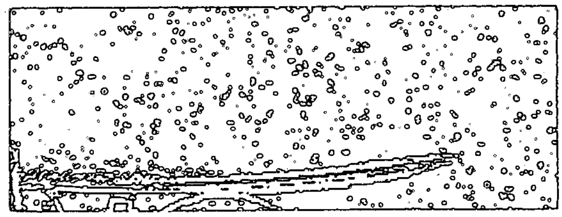
The 3-D views and the binary edges of the test images are shown in Fig. 7. The edge images are taken using Sobel operator with threshold 30 and used as input measurements of data association system. Notice that the boundaries are very smooth, thick and very noisy.



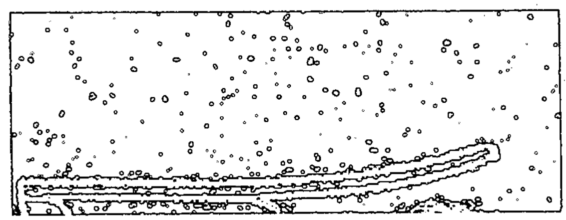
(a)



(b)



(c)



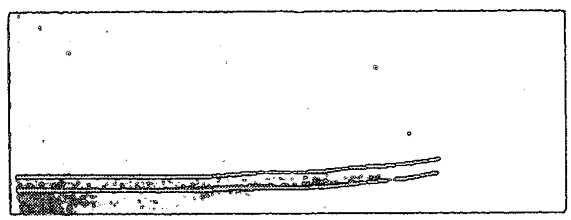
(d)

Fig. 7. (a) and (b) 3-D views of the test images, and (c) and (d) edge images (sobel with threshold 30).

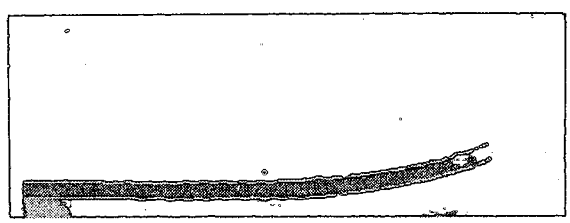
The results are shown in Fig. 8.

The upper and lower boundaries obtained by the proposed method are shown with the original images.

Notice, the upper and lower boundaries are correctly detected.



(a)



(b)

Fig. 8. (a) and (b) upper and lower boundaries of the test images.

VII. Conclusion

In this paper we introduced an algorithm for object boundary detection that consists of data association and tracking filter. We applied the basic concept of data association technique proposed in the field of multiple target tracking to detection of object boundary. After defining the data association as a constrained optimization, we introduced a new energy function and thereby a Hopfield neural network as an efficient method for solving the energy function. The new algorithm was tested for finding object boundaries in IR images.

References

[1] T. E. Fortmann, Y. Bar-Shalom, *Tracking and Data Association*, Orland Acdemic Press, Inc., pp. 224, 1988.

[2] J. E. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Royal Statics.* Soc., Ser. B 36, pp. 192-236, 1974.

[3] T. E. Fortmann, Y. Bar-Shalom and M. Scheffe, "Sonar tracking of multiple targets using joint probability data association," *IEEE Journal of Oceanic Engineering*, Vol. OE-8, No. 3, pp. 173-183, July, 1983.

[4] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley Pub., Inc., 1992.

[5] J. J. Hopfield and D. W. Tank, "Neural computation of decisions in optimization problems," *Biological Cybernetics*, 52, pp. 141-152, 1985.

[6] J. Kim and H. Jeong, "Neural networks for optimal data association of hot plate," in Proc. of WCNN'95, INNS Press, 1995.

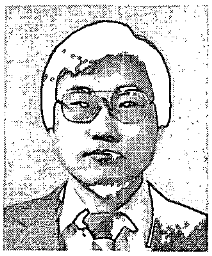
[7] Y. Lee and H. Jeong, "A neural network approach to the optimal data association in multi-target tracking," in Proc. of WCNN'95, INNS Press, 1995.

[8] D. Sengupta and R. A. Iltis, "Neural solution to the multitarget tracking data association problem," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. AES-25, No. 1, pp. 96-108, Jan. 1989.

[9] P. Smith and G. Buechler, "A branching algorithm for discrimining and tracking multiple objects," *IEEE Trans. on Automatic Control*, Vol. AC-20, pp. 101-104, Feb. 1975.

[10] H. M. Sun and S. M. Chiang, "Tracking multitarget in cluttered environment," *IEEE Trans. on Aerospace and Electronic System*, Vol. AES-28, No. 2, pp. 546-559, April, 1992.

Jung-Gu Kim was born in Jin-Ju Kyoung-Buk Korea, in 1964. He received the B.S. degree in the Department of Electrical Engineering from the Kyoung-Buk National University in 1991, and the M.S. degree in the Department of Electrical Engineering from the Pohang University of Science and Technology(POSTECH) in 1993. Since 1993, he has been working towards the Ph.D. degree in the Electrical Engineering from POSTECH. His research interests are computer vision and digital image processing.



Hong Jeong was born in Seoul, Korea, in 1953. He received B.S. degree in the Department if Electrical Engineering from the Seoul National University in 1977. In 1979, he received the M.S. degree in the Department of Electrical Engineering from the Korea Advanced Institute of Science and Technology. In 1984, 1986, and 1988, he received the S.N., E.E., and Ph.D. degree, respectively, all in the Department of Electrical Engineering and Computer Science at M.I.T., Cambridge, massachusetts, U.S.A. During the period of 1979-1982, he was a faculty staff at the Department of Electrical Engineering at the Kyoung-Buk National University, Daegu, Korea. Since 1988, he was worked in the Department of Electrical Engineering at the Pohang University of Science and Technology, where he now works as a Professor. He is a Sigma Xi member. During 1994-1995, he worked as a vice-chairman in the Special Interest Group on Neurocomputing in the Korea Information Science Society. Also from 1991, he has worked as a committe staff in the Neural Networks, Fuzzy, and Artificial Intelligence Group in the Korean Institute of Telematics and Electronics. His research interests include digital signal processing, computer vision, speech recognition, and radar signal processing.