

Pattern Recognition 35 (2002) 2031-2040

PATTERN RECOGNITION THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Real-time tracking of multiple objects in space-variant vision based on magnocellular visual pathway ☆

Seonghoon Kang, Seong-Whan Lee*

Department of Computer Science and Engineering/Center for Artificial Vision Research, Korea University, Anam-dong, Seongbuk-ku, Seoul 136-701, South Korea

Received 7 March 2001; received in revised form 2 October 2001

Abstract

In this paper, we propose a space-variant image representation model based on properties of magnocellular visual pathway, which perform motion analysis, in human retina. Then, we present an algorithm for the tracking of multiple objects in the proposed space-variant model. The proposed space-variant model has two effective image representations for object recognition and motion analysis, respectively. Each image representation is based on properties of two types of ganglion cell, which are the beginning of two basic visual pathways; one is parvocellular and the other is magnocellular. Through this model, we can get the efficient data reduction capability with no great loss of important information. And, the proposed multiple objects tracking method is restricted in space-variant image. Typically, an object-tracking algorithm consists of several processes such as detection, prediction, matching, and updating. In particular, the matching process plays an important role in multiple objects tracking. In traditional vision, the matching process is simple when the target objects are rigid. In space-variant vision, however, it is very complicated although the target is rigid, because there may be deformation of an object region in the space-variant coordinate system when the target moves to another position. Therefore, we propose a deformation formula in order to solve the matching problem in space-variant vision. By solving this problem, we can efficiently implement multiple objects tracking in space-variant vision.

Keywords: Space-variant vision; Biologically motivated vision; Multiple objects tracking

1. Introduction

In developing an active vision system, there are three main requirements imposed on the system: high resolution for obtaining details about the regions of interest, a wide field of view for easy detection of a looming object or an interesting point, and the fast response time of the system [2]. However, a system that uses a traditional image representation which has uniformly distributed resolution cannot satisfy such requirements. Because the image size must be very large in order to include details of interest region and whole environment, it is impossible to process whole image in real-time. Therefore, many methods to solve these problems have been developed. Some methods use wide-angle camera with an additional zoom camera to process a wide range of environments and detailed interest region. But these methods are not ultimate solutions for these problems.

There have been many research works on effective image representation models based on the biological vision system that satisfies all of the requirements, among which is the

[☆] The preliminary version of this paper has been presented at the First IEEE International Workshop on Biologically Motivated Computer Vision, Seoul, Korea, in May 2000 [1]. This research was supported by Creative Research Initiatives of the Ministry of Science and Technology, Korea.

^{*}Corresponding author. Tel.: +82-2-3290-3197; fax: +82-2-926-2168.

E-mail addresses: shkang@image.korea.ac.kr (S. Kang), swlee@image.korea.ac.kr (S.-W. Lee).

space-variant model using the multi-resolution property of the biological vision system.

Animals including humans do not transfer whole input information but only a few data from eyes to the brain by data reduction (about 128:1) due to structural properties of retina. For example, in monkey's retina, there is a central region called "fovea" in which the photoreceptors are most densely packed, and the photoreceptor's density becomes smaller, decaying towards the outer periphery. The log(z)model [3], which is the first model considering such properties of biological vision system, was developed. And many variations of this model have been investigated. But, these models have considered only a few properties (many important properties have been ignored). Thus, in this work, the new space-variant image representation model is proposed, which considers some important properties that were not considered in previous models.

Such space-variant models have recently been applied to many active vision applications such as moving object tracking, vergence control, etc. In a typical active tracking system, there is only one target to track because the camera head continuously fixates on one object at a time. In order to transfer its fixation to another object, it is important to keep track of the positions of the objects moving in the background. This gives rise to the necessity for multiple objects tracking. In multiple objects tracking, new problems are introduced, such as multiple objects detection, and matching between the current target object and the detected object in the space-variant coordinate system.

In space-variant vision, multiple objects detection is difficult to achieve, because a motion vector in the space-variant coordinate system is represented differently from that of the Cartesian one in size and direction. Consequently, it is difficult to segment a moving region directly in the space-variant coordinate system. The matching problem is also very difficult, because there may be deformation of an object region in the space-variant coordinate system when it moves to another position, although the target object is rigid. In this paper, we propose an efficient algorithm that overcomes the difficulties mentioned above.

2. Related works

Most of the space-variant image representation models are based on Schwartz's log(z) model. These models are divided into two models, one is conformal mapping, and the other is overlapped mapping.

First, we consider a conformal mapping. A conformal mapping is a function of a complex variable that has the property of preserving relative angles. There is Schwartz's $\log(z)$ model [3] which is the representative of space-variant model. This model uses complex-logarithmic function, $w = \log(z)$. It has size and rotation invariant properties. But it needs a uniformly distributed resolution patch at the center of the image as shown in Fig. 1, because it has singularity at



Fig. 1. Schwartz's log(z) model.



Fig. 2. Bolduc's overlapped mapping model.

origin due to characteristic of complex-logarithmic mapping function.

We also consider overlapped mapping models. Most of the space-variant models belong to this category. The most representative model is Bolduc's model [2]. It is very similar to Schwartz's model, but, as shown in Fig. 2, its receptive fields are circular in form and adjacent receptive fields overlap each other. Therefore it is more similar to biological vision system.

Another overlapped mapping model is that of Wilson [4]. This model is very similar to Bolduc's model, but it uses log(z + a) mapping function. This mapping function needs no uniform resolution patch image, solving the singularity problem in log(z) model. But there is a disadvantage of discontinuity between the left and right planes of an image.

Recently, Yamamoto's model [5] was proposed. This model is different from other overlapped models, because of the existence of photo-receptor layer in front of transform layer.

All of the above models use the simplified property of biological vision system, considering only the fact that the photoreceptor's density is high in the foveal area and low in the peripheral area. These models do not consider the separated visual pathways. Thus, we propose a more efficient and biologically plausible model by including more properties of biological vision system.

These space-variant models have been applied to various areas, such as active vision, visual perception, etc. Panerai et al. [6] developed a technique for vergence and tracking in log-polar images. Lim et al. [7] proposed a tracking algorithm for space-variant active vision. Recently, Jurie [8] proposed a new log-polar mapping procedure for face detection and tracking.

There have also been many works on multiple objects tracking. Recently, Bremond and Thonnat [9] presented a method to track multiple non-rigid objects in a video sequence, and Haritaoglu et al. [10] developed Hydra, a real-time system for detecting and tracking multiple people.

3. Multiresolutional properties of biological vision system

The human retina consists of three layers constructed from six types of cells (photoreceptor, horizontal cell, amarcrine cell, interplexiform cell, bipolar cell, ganglion cell). The light is translated into the neural signal at photoreceptor. Then, this signal is transferred to the brain through the many cells in the retina [11]. We remember that the number of photoreceptors, which transform light into neural signal, is greater than the number of ganglion cells, which are the final gate of neural signal's transferring path in retina (the ratio is about 128:1). According to above fact, we can say that a few data are transferred to the brain, after data are reduced at retina. Thus, in this chapter, we investigate the data reduction property of retina through physiological evidence.

3.1. Distribution property of photoreceptors

Photoreceptors consist of cone and rod cells. Cone cells, which respond to color and bright light, are more important than rod cells. These cone cells are densely packed in the fovea, and their density rapidly decays towards the outer periphery. This property is shown in Fig. 3 [12]. Most of the previous space-variant models are based on this distribution property of photoreceptors.

3.2. Properties of ganglion cells

Ganglion cells are also densely packed in the foveal region and their density decays towards the outer periphery. As shown in Fig. 4, it is well known that the decay rate is inversely proportional to the square of eccentricity [13].

There are two types of ganglion cells, one is P cell (Parvocellular cell) and the other is M cell (Magnocellular cell). The former is the beginning of the recognition path and the latter is the beginning of the motion analysis path. These cells have different properties. The main difference is the size of its receptive field. As shown in Fig. 5, there is a specific relation between the size of receptive field and eccentricity. Watson [14] explained this relation with the following, so called, scaling function, where e and s are eccentricity and scaling factor, respectively.

$$s = 1 + ke. \tag{1}$$

Suppose that the size of receptive field in the foveal region is *w*, the size of receptive field at eccentricity *e* is *sw*.

In this paper, we use this scaling function in order to design a space-variant image representation model, and we



Fig. 3. Distribution of photoreceptors by eccentricity.



Fig. 4. Distribution of ganglion cells by eccentricity.



Fig. 5. Size of receptive field by eccentricity.

use an approximated value from physiological experiments data for constant k.

4. Proposed space-variant representation model

Previous well-known space-variant models are based on simplified properties of ganglion cell where its receptive field size is small in the fovea region and larger towards the outer periphery. But there are two types of ganglion cells (M and P cells), which have different purpose of process and receptive field size. Thus, in this paper, we design an effective model better than previous simplified models, by overcoming their limitations.



Fig. 6. Receptive fields in the peripheral region.

4.1. Modeling of receptive field

The size of receptive field is determined by scaling function, Eq. (1). We model the size of receptive field using this scaling function, and we determine the scaling function by analysis of data from physiological experiments [15] about receptive field of M and P cells. First, we get the approximated scaling functions in Eqs. (2) and (3), where the subscript letters M and P mean M cell and P cell, respectively (specially, from the data from Perry's physiological experiments, $k_{\rm M}$ is about 0.44 and $k_{\rm P}$ is about 0.78).

$$s_{\rm M} = 1 + k_{\rm M} e, \tag{2}$$

$$s_{\rm P} = 1 + k_{\rm P} e. \tag{3}$$

Suppose that the receptive field size of M and P cells in the foveal region is w_{M0} and w_{P0} , respectively. The relations between the size of receptive field and eccentricity are given in Eqs. (4) and (5) (specially, from the data of Perry's physiological experiments, w_{M0} is about 0.07° and w_{P0} is about 0.01°).

$$w_{\rm M} = w_{\rm M0} + w_{\rm M0} k_{\rm M} e, \tag{4}$$

$$w_{\rm P} = w_{\rm P0} + w_{\rm P0} k_{\rm P} e. \tag{5}$$

These relations of M and P cells are the same. Thus, we use the same model for M and P cells, and only parameters of its model are different. And, we consider the model for the foveal region and the peripheral region separately, because the distribution of receptive field is uniform in the foveal region, and is more sparse towards the outer periphery.

4.1.1. Peripheral region

As shown in Fig. 6, the size of receptive field is larger towards the outer periphery. Eq. (6) is the relation between the eccentricity of the (n - 1)th ring and the *n*th ring. Substituting Eq. (7) which means the receptive field size by eccentricity for w_n in Eq. (6), then we get Eq. (8) about eccentricity of the *n*th ring, where R_n is the eccentricity of the *n*th ring, w_0 is the receptive field size at the *n*th ring, w_0



Fig. 7. Receptive fields in the foveal region.

is the receptive fields size at the foveal region, k is constant and o is the overlapping factor. When the overlapping factor is one, it means that receptive fields are completely overlapped, and when it is zero, it means that receptive fields are not overlapped.

$$R_{n-1} + \frac{w_{n-1}}{2} - ow_n + \frac{w_n}{2} = R_n,$$
(6)

$$w_n = w_0 + w_0 k R_n, \tag{7}$$

$$R_n = \frac{2(1-o)w_0 + (2+w_0k)R_{n-1}}{2-(1-2o)w_0k}.$$
(8)

Because the number of receptive fields at each ring is the same, Eq. (9) is established, where round function means the rounding off operation, and R_0 is the size of the foveal region.

$$K = round\left(\frac{2\pi R_0}{(1-o)(w_0 + w_0 k R_0)}\right).$$
(9)

4.1.2. Foveal region

As shown in Fig. 7, the size of receptive field is the same in the foveal region regardless of eccentricity. Therefore the number of receptive fields at each ring varies according to eccentricity. This relation is shown in Eq. (10). Also, we can generalize the function about eccentricity of each ring to Eq. (11), considering the size of receptive field and the overlapping factor.

$$R_n = R_{n-1} - \frac{R_0}{round(R_0/(1-o)/w_0)},$$
(10)

$$K_n = round\left(\frac{2\pi R_n}{(1-o)w_0}\right).$$
(11)

4.2. Template of space-variant representation

We can compose the mapping templates which transform a traditional image to a space-variant image using the constructed model. We compose two types of mapping template which are called M- and P-Map. Each template is based on properties of M-cell and P-cell, respectively. We can obtain two types of image; one is used for object recognition and the other is for motion analysis. Given an image of 512×512 pixels having a viewing angle of 100° , and an overlapping factor of 0.3, we get the templates as in Figs. 8 and 9.

5. Deformation formula in space-variant vision

There may be a deformed movement in the space-variant coordinate system, which corresponds to a simple translation in the Cartesian coordinate system, as shown in Fig. 10. We should be aware of the deformation formula in order to analyze the movement of a region in the space-variant coordinate system.

In Fig. 11, we know the deformed position of a certain point in the space-variant coordinate system that corresponds to the translated position in the Cartesian coordinate system. The radius *R* and the angle θ of the given point (ξ, η) in the space-variant coordinate system can be obtained using the mapping functions (6) and (7) given below

$$R = SM_R(\xi) = \sum_{n=1}^{\xi} a b^{n-1} + R_0 b^{\xi}$$
$$= \frac{(a+bR_0 - R_0)b^{\xi} - a}{b-1},$$
(12)

$$\theta = SM_{\theta}(\eta) = \frac{2\pi}{K}\eta, \tag{13}$$

where $a = 2(1 - o)w_0/(2 - (1 - o)w_0k)$ and $(2 + w_0k)/(2 - (1 - 2o)w_0k)$. Then, the position (R', θ') of the point in the Cartesian coordinate system, after the movement of the point by Δx and Δy , can be found easily, as follows:

$$R' = \sqrt{(x + \Delta x)^2 + (y + \Delta y)^2},$$
(14)

$$\theta' = \arctan\left(\frac{y + \Delta y}{x + \Delta x}\right),$$
(15)

where $x = R \cos \theta$, $y = R \sin \theta$. From the equations given above, the deformed position(ξ' , η') in the space-variant coordinate system can be derived as

$$\xi' = SM_R^{-1}(R') = \log_b \left(\frac{(b-1)R' + a}{a + bR_0 - R_0} \right),$$
(16)

$$\eta' = SM_{\theta}^{-1}(\theta') = \frac{K}{2\pi}\theta'.$$
(17)

Finally, the deviation $(\Delta \xi, \Delta \eta)$ in the space-variant coordinate system can be obtained as follows:

$$\Delta \xi = SM_R^{-1}(R') - \xi, \tag{18}$$

$$\Delta \eta = SM_{\theta}R^{-1}(\theta') - \eta.$$
⁽¹⁹⁾



Fig. 8. Mapping template of M-Map.



Fig. 9. Mapping template of P-Map.



Fig. 10. (a) A movement in the Cartesian coordinate system, and (b) the deformed movement in the space-variant coordinate system, which corresponds to (a).



Fig. 11. Deformed movement in the space-variant coordinate corresponding to movement in the Cartesian coordinate.

As we have shown above, the deformed deviation $(\Delta \xi, \Delta \eta)$ of a point (ξ, η) , which comes from a translation $(\Delta x, \Delta y)$ in the Cartesian coordinate system, can be found when the point (ξ, η) and the translation $(\Delta x, \Delta y)$ are given.

6. Moving object detection

6.1. Motion estimation in space-variant vision

Horn and Schunk's optical flow method [16] is used for motion estimation. In addition, the optical flow vectors in the space-variant coordinate system are transformed to the Cartesian coordinate system in order to segment a moving region and to calculate a mean flow vector easily, as shown in Fig. 12.

A vector $\mathbf{A}_{S} = A_{\xi} \mathbf{a}_{\xi} + A_{\eta} \mathbf{a}_{\eta}$ in the space-variant coordinate system can be represented as a polar coordinate vector $\mathbf{A}_{S} = A_{\xi}(R_{n} - R_{n-1})\mathbf{a}_{\rho} + A_{\eta}K_{\eta}\mathbf{a}_{\phi}$. Then, it is easily transformed to a Cartesian coordinate vector $\mathbf{A}_{C} = A_{x}\mathbf{a}_{x} + A_{y}\mathbf{a}_{y}$ where

$$A_{x} = \mathbf{A}_{S} \cdot \mathbf{a}_{x}$$

= $A_{\xi}(R_{n} - R_{n-1})\mathbf{a}_{\rho} \cdot \mathbf{a}_{x} + A_{\eta}K_{\eta}\mathbf{a}_{\phi} \cdot \mathbf{a}_{x}$
= $A_{\xi}(R_{n} - R_{n-1})\cos(K_{\eta}\eta) - A_{\eta}K_{\eta}\sin(K_{\eta}\eta),$ (20)

$$\begin{aligned} A_{y} &= \mathbf{A}_{S} \cdot \mathbf{a}_{y} \\ &= A_{\xi} (R_{n} - R_{n-1}) \mathbf{a}_{\rho} \cdot \mathbf{a}_{y} + A_{\eta} K_{\eta} \mathbf{a}_{\phi} \cdot \mathbf{a}_{y} \\ &= A_{r} (R_{n} - R_{n-1}) \sin(K_{\eta} \eta) + A_{\eta} K_{\eta} \cos(K_{\eta} \eta). \end{aligned}$$
(21)

6.2. Motion estimation in space-variant vision

For moving region segmentation, a region-based segmentation and labeling are employed. As shown in Fig. 13 we construct a binary optical flow map, then morphological post-filtering is applied. Finally, a labeled region map is obtained by using a connected component analysis.



Fig. 12. Vector transformation from the space-variant coordinate system to the Cartesian coordinate system.



Fig. 13. Region-based moving region segmentation.

7. Multiple objects tracking

7.1. Motion estimation in space-variant vision

For multiple objects tracking, we construct a dynamic target model which consists of the texture of an object region. This target model is updated continuously in time during the tracking process. It is defined by

$$\Psi^{t}(\xi,\eta) = wI(\xi,\eta) + (1-w)D_{\Psi^{t-1},\Delta P}(\xi,\eta),$$
(22)

where $\Psi^{t}(\xi,\eta)$ is the intensity value at (ξ,η) in a texture template, $I(\xi,\eta)$ is that of the input image, $D_{\Psi^{t-1},\Delta P}(\xi,\eta)$ is that of a deformed template at (ξ,η) and w(0 < w < 1) is a weight value.

7.2. Deformable matching of moving targets

For matching between the target model and a detected object, we use a correlation technique, the quadratic difference, defined by

$$C_{ij} = \frac{\sum_{\xi,\eta\in\Psi} (I(\xi+i,\eta+j) - D_{\Psi,\Delta P}(\xi,\eta))^2}{N^2},$$
(23)

where N is the number of pixels in the target template Ψ and i, j < W, W is the size of a search window. When C_{ij} is smaller than a certain threshold value, the detected object is matched with the current target. The target model is then updated to the detected object (Fig. 14).



Fig. 14. Examples of extracted target model.

Now, we can track multiple objects by using the detection, prediction, matching, and updating loop.

8. Experimental results and analysis

Our experimental system consists of a Pentium III 500 MHz PC with a Matrox Meteor II image grabber, and Cannon VC-C1 CCD camera.

The performance of multiple objects tracking in space-variant active vision is shown. In our experiment, people moving around in an indoor environment are chosen as target objects to track. During the experiment, they are walking across the field of view in various situations.

Fig. 15 shows the result images of each sub-process of the tracking process. The images in the third row are binary optical flow maps, each of which is constructed by transformation of the directions of the optical flow vectors into a gray-level image with an appropriate threshold. The images in the forth row are morphological post-filtered maps, each of which is constructed by a morphological closing operation. Each of these maps is used as an extracting mask for target model construction. Finally, the images in the last row are segmented target models. The target model is used for matching between the tracked objects in the previous frame and the detected objects in the current frame. Throughout these processes, multiple objects tracking is accomplished.



Fig. 15. Result images of sub-processes of our system: (a) is a Cartesian image; (b) is a space-variant image; (c) is an optical flow map; (d) is a morphological post-filtered map; (e) is an extracted target region and (f) is an extracted target model.



Fig. 16. Tracking sequence: images on odd row are tracking sequences and images on even row are reconstructed images from detected target region.

Fig. 16 shows a sequence of images taken by the CCD camera. As shown in the figure, the system did not lose the targets. In the previous moving object tracking algorithms using traditional image representation, it is difficult to detect and track moving objects with small motion in real-time. So, in oder to implement real-time system, the system could detect and track moving object roughly, because it must process input image in coarse level. But, the proposed algorithm can detect and track moving objects with small motion in the region of interest; also, it can be processed in real time. However, it detected only one target for two objects upon occlusion, since occlusion is not considered in our system. Future research will take this into account for a better tracking performance.

9. Conclusions and further researches

In this paper, we have shown multiple targets tracking in space-variant active vision in a low-cost PC environment. Motion detection is very efficient because of using optical flows in the space-variant coordinate system. Considering deformation in the space-variant coordinate system, caused by the movement of the target region, matching in space-variant vision becomes very efficient and simple. Nevertheless, the proposed algorithm does not consider occlusion problems–solutions which should be included for better tracking performance. Also, it has the drawback that it cannot detect the motion of small objects in far periphery,



Fig. 17. Motion of small object in far periphery.

as shown in Fig. 17. In this case, we should solve this problem by enhancing the processing rate by selective attention technique.

References

- S. Kang, S-W. Lee, Multiple object tracking in multiresolution image sequences, Proceedings of First IEEE International Workshop on Biologically Motivated Computer Vision, Seoul, Korea, May 2000, pp. 564–573.
- [2] M. Bolduc, M.D. Levin, A real-time foveated sensor with overlapping receptive fields, Real-Time Imaging 3 (3) (1997) 195–212.
- [3] E.L. Schwartz, Spatial mapping in the primate sensory projection: analytic structure and relevance to perception, Biol. Cybern. 25 (1977) 181–194.
- [4] S.W. Wilson, On the retino-cortical mapping, Int. J. Man-Machine Studies 18 (1983) 361–389.
- [5] H. Yamamoto, Y. Yeshurun, An active foveated vision system: attentional mechanisms and scan path convergence measures, Comput. Vision Image Understanding 63 (1) (1997) 50–65.
- [6] F. Panerai et al., Space variant vision for an active camera mount, Technical Report TR1/95, LIRA-Lab-DIST University of Genova, 1995.
- [7] F.L. Lim et al., Tracking in a space variant active vision system, Proceedings of 13th International Conference on Pattern Recognition, Vienna, Austria, 1996, pp. 745–749.
- [8] F. Jurie, A new log-polar mapping for space variant imaging: application to face detection and tracking, Pattern Recognition 32 (5) (1999) 865–875.
- [9] F. Bremond, M. Thonnat, Tracking multiple nonrigid objects in video sequences, IEEE Trans. Circuits Systems Video Technol. 8 (5) (1998) 585–591.
- [10] I. Haritaoglu et al., Hydra: multiple people detection and tracking using silhouettes, Proceedings of Second IEEE Workshop on Visual Surveillance, Fort Collins, Colorado, 1999, pp. 6–13.
- [11] D.H. Hubel, Eye, Brain and Vision, Scientific American Library, 1995.
- [12] M.H. Pirenne, Vision and the Eye, 2nd Edition, Chapman & Hall, London, UK, 1967.
- [13] R.E. Kronauer, Y.Y. Zeevi, Reorganization and diversification of signals in vision, IEEE Trans. System, Man Cybern. 15 (1985) 91–101.
- [14] A.B. Watson, Detection and recognition of simple spatial forms, Phys. Biol. Process. Images (1982) 100–114.

- [15] V.H. Perry et al., Retinal ganglion cells that project to the dorsal LGN in Macaque monkeys, Neuroscience 12 (1984) 1101–1123.
- [16] A.M. Tekalp, Digital Video Processing, Prentice-Hall, Englewood Cliffs, NJ, 1995.

About the Author—SEONGHOON KANG was born in Seoul, Korea, in 1973. He received the BS and MS degrees in Control and Instrumentation Engineering from Korea University, Seoul, Korea, in 1995 and 1997. He is a Ph.D. candidate in the Department of Computer Science and Engineering, Korea University, Seoul, Korea.

His research interests include motion and object detection, active vision and mobile visual aids.

About the Author-SEONG-WHAN LEE received his BS degree in Computer Science and Statistics from Seoul National University, Seoul, Korea, in 1984, the MS and Ph.D. degrees in computer science from KAIST in 1986 and 1989, respectively. From February 1989 to February 1995, he was an assistant professor in the Department of Computer Science at Chungbuk National University, Cheongju, Korea. In March 1995, he joined the faculty of the Department of Computer Science and Engineering at Korea University, Seoul, Korea, as an associate professor, and now he is a full professor. Currently, Dr. Lee is the director of National Creative Research Initiative Center for Artificial Vision Research (CAVR) supported by the Korean Ministry of Science and Technology. He was the winner of the Annual Best Paper Award of the Korea Information Science Society in 1986. He obtained the First Outstanding Young Researcher Award at the Second International Conference on Document Analysis and Recognition in 1993, and the First Distinguished Research Professor Award from Chungbuk National University in 1994. He also obtained the Outstanding Research Award from the Korea Information Science Society in 1996. He has been the co-Editor-in-chief of the International Journal on Document Analysis and Recognition since 1998 and the associate editor of the Pattern Recognition Journal, the International Journal of Pattern Recognition and Artificial Intelligence, and the International Journal of Computer Processing of Oriental Languages since 1997. He was the Program co-chair of the Sixth International Workshop on Frontiers in Handwriting Recognition, the Second International Conference on Multimodal Interface, the 17th International Conference on Computer Processing of Oriental Languages, the Fifth International Conference on Document Analysis and Recognition, and the Seventh International Conference on Neural Information Processing. He was the workshop co-chair of the Third International Workshop on Document Analysis Systems and the First IEEE International Workshop on Biologically Motivated Computer Vision. He served on the program committees of several well-known international conferences. He is a fellow of IAPR, a senior member of the IEEE Computer Society and a life member of the Korea Information Science Society and the Oriental Languages Computer Society. His research interests include pattern recognition, computer vision and neural networks. He has more than 150 publications on these areas in international journals and conference proceedings, and has authored five books.