Information Processing in Primate Retinal Ganglion

Sung-Kwan Je, Jae-Hyun Cho and Gwang-Baek Kim, *Member, KIMICS*

Abstract—Most of the current computer vision theories are based on hypotheses that are difficult to apply to the real world, and they simply imitate a coarse form of the human visual system. As a result, they have not been showing satisfying results. In the human visual system, there is a mechanism that processes information due to memory degradation with time and limited storage space. Starting from research on the human visual system, this study analyzes a mechanism that processes input information when information is transferred from the retina to ganglion cells. In this study, a model for the characteristics of ganglion cells in the retina is proposed after considering the structure of the retina and the efficiency of storage space. The MNIST database of handwritten letters is used as data for this research, and ART2 and SOM as recognizers. The results of this study show that the proposed recognition model is not much different from the general recognition model in terms of recognition rate, but the efficiency of storage space can be improved by constructing a mechanism that processes input information.

Index Terms—Artificial vision, Compression, Image recognition, artificial neural network, Retinal Ganglion

I. INTRODUCTION

Along with the development of computers, a lot of efforts to realize artificial intelligence through recognizable computers have been continued. Research on the human visual system has been trying to improve the limitations of computers, which cannot respond to various environmental changes. Ongoing vision studies, however, imitate part of the function of human visual system and develop machine vision systems that can be supplied to machines. Because of the limited application of machine vision technology to the real world environment, information processing of machine vision systems is still quite different from that of human visual system.

Recently, many developed countries are actively conducting research to maximize the performance of computer vision technology and to develop artificial vision through the modeling of human visual processing [3, 4, 22, 23]. Artificial vision is to develop information processing procedures of the human visual system based on the biological characteristics. Compared with the machine vision technology, it can be effectively applied to industry. By investing over 20 billion yen between 1997 and 2016, Japan is conducting research on the areas of machine intelligence, voice recognition and artificial vision based on the information processing mechanism of the brain. Centered around the National Science Foundation (NSF) and the Application of Neural Networks for Industries in Europe (ANNIE), America and Europe are also conducting research on artificial vision, as well as artificial intelligence and voice recognition using the modeling of the brain's information processing [3,4].

Machine vision systems process all images regardless of objects and background. In recognizing objects, it exhausts calculation expenses in the images of unnecessary background. By storing unnecessary background images without processing, the efficiency of storage space is lowered. Inefficiency of storage space can influence the maintenance of the system. Therefore, it is necessary to have a mechanism that can process input information like the human visual system.

This study is to propose a model that can implement a mechanism of information processing that is actually occurring in the human visual system. We will review information processing procedures of the retina in Chapter 2, construct a model based on the mechanism in Chapter 3, and present the results of the experiment and conclusions in Chapter 4.

II. INFORMATION PROCESSING IN RETINA

The processing of human visual information is composed of several stages. The first stage of information processing occurs in the retina. The retina not only converts light energy to electrochemical energy, but transmits the information to the visual path.

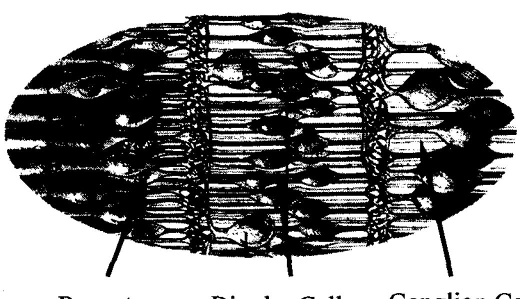


Fig. 1 Retinal Cells [22, 23]

Manuscript received May 15, 2004.

S. K. Je is with the Department of Computer Sciences, Pusan National University, Busan, Korea. (e-mail:jimmy374@pusan.ac.kr)

J. H. Cho is with the Department of Computer Information, Catholic University of Pusan, Busan, Korea. (corresponding author to provide e-mail:jhcho@cup.ac.kr)

G. B. Kim is with the Department of Computer Engineering, Silla University, Busan, Korea. (e-mail:gbkim@silla.ac.kr)

A human visual system describes in figure1. As for the human visual system, a light energy becomes a shadow with a reversed phase in the retina, and it transmits a displayed neural signal to ganglion cell and, as for this image, it is through a cognition process on delivery with primary visual cortex through optic nerve later by the retina. The retina located behind a pupil is complicatedly composed of ten vertical layers. However, since ganglion cell is composed with only about 1,000,000 in one eye, a lot of visual information is compressed for transferring from 125,000,000 receptors to ganglion cell. Actually it is turned into a neural signal by an operation of the retina, but the sampling image is transmitted to primary visual cortex of a brain through optic nerve of the retina which listened to this signal in 125,000,000 rod cells and 6,000,000 con cells.

Also, one fovea sampling of non-uniform is performed in receptor level. This means that sampling is dense in a fovea region, but loose in peripheral region because output of a con cell is transmitted from peripheral region to ganglion cell in fovea by more limited dynamic range. This process means the important data compression occurs [4]. It is the place where a process is originated conversion of visual information becomes quit at the same time because the retina has both a receptor and neuron.

Ganglion cells can be mapped into P-cells and M-cells. P-cells contain major information of images 'what', whereas M-cells contain edge information of images 'where'[4]. The information is decomposed into i) processes to recognize objects and ii) processes to recognize the location of objects. It depends on a mechanism that can minimize the loss of critical information when transmitting information and can maximally minimize the amount of total information. That is, information related to perceiving 'What' is transmitted to P-cells; and P-cells comprises 80% of total ganglion cells and minimize the loss during transmission. Whereas, information related to 'Where' is sent to M-cells; and M-cells comprises 20% of total ganglion cells [22, 23]. Therefore, human vision is sensitive to 'What' information, but insensitive to 'Where' information. It values information about objects, but compresses information about background maximally. Even if a lot of information is compressed, human beings do not displaying a particular problem in recognizing images [5-6].

Machine vision systems, however, recognize images without this processing, and exhaust calculation expenses to recognize objects along with unnecessary background images. In addition, because they store unnecessary background images without any processing, the efficiency of storage space is also reduced. This inefficiency of storage space can influence greatly the maintenance of systems. Therefore, a mechanism like the human visual system is needed to process input information.

Image compression technology can provide clear images even at high compression rate, and it considers important characteristics of images. Discrete cosine transform (DCT) that has been used for existing compression techniques transforms images by cutting them into regular sizes of blocks. As compression rates are increased, a blocking phenomenon occurs; as a result, the quality of images is decreased drastically. To overcome these shortcomings, the lapped orthogonal transform (LOT), which uses overlapping of each block, and the wavelet transform have been proposed. Especially, the wavelet transform, which performs frequency and resolution decomposition, is quite different from the existing block transforms; it focuses on removing the relativity among vectors while keeping the relativity within the block, they are suitable for dealing with important information (what) of images. The wavelet transform has been adopted as JPEG2000 and has the following characteristics [11].

a) Large image representation without a blocking phenomenon
b) Excellent compression at low bit rate
c) Loss and lossless compression are possible with one algorithm
d) The progressive transmission by resolution
e) Real-time coding

Using the characteristics of the human visual system, this study implement a mechanism that processes input information when transmitting information from the retina to ganglion cells, and compares it to existing visual models and their performance.

III. PROPOSED ALGORITHM

A. Information Processing in Retinal Ganglion

According to the current computer vision theory, the original image received goes though extracting feature of an object and a process of recognizing an object. A conventional computer vision process is same as figure 2. That is, it imitates human visual characteristics of extracting information in computer vision and is recognizing an object. However, the algorithm of current computer has been developing as a different form from human visual information processing in the actual application. Study about information processing of a human brain and a visual information processing are actively attempted recently. And a human visual information processing is gradually figured out. Currently study of these models is attempted actively each nation in the world [1-4]. Computer vision is emulating a rough form of human visual information processing, and the information processing is showing them the form that is different from a human information processing [22, 23].

Due to this problem, machine visual systems cannot adapt to the changes of the environment, and have a lot of problems in real world application.

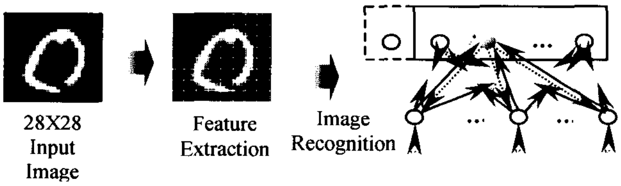


Fig. 2 General Recognition Process of Computer Vision

In this paper, a model is proposed to implement a mechanism that processes visual information, occurring in the recognition process of the real human visual system. A flowchart of the general algorithm is presented in Figure 3. In the retina, information is transmitted by minimizing the loss of important information in ganglion cells and by executing a mechanism that minimizes total amount of information maximally.

For example, the reason that we have a good impression of a cafe is because we remember the interior (pictures on the walls, the shape and the location of tables, flowers on the table), outfit of employees, cafe owner, and music of the cafe, and reflect on the overall ambience of the cafe. When we talk about an impression of this cafe to a third person, we explain these objects from our memory.

band has very much delicate information to express original image in wavelet transform and much information of original image exists. As for theses, a mapping process of parvo-cellular of ganglion cell (P-cell) and the mechanism are very alike in receptor of the retina. Also, information of high-band of wavelet transform has edge information of original image compared to low-band which has a little information. These mechanisms are very similar to a mapping process of magno-cellular of ganglion cell (M-cell). Therefore, P-cell of ganglion cell deals with main information of an image like low-band, and M-cell is dealing with edge information of an image like the high-band. In this paper, we used wavelet transform in a compression process of this visual information and composed the model.

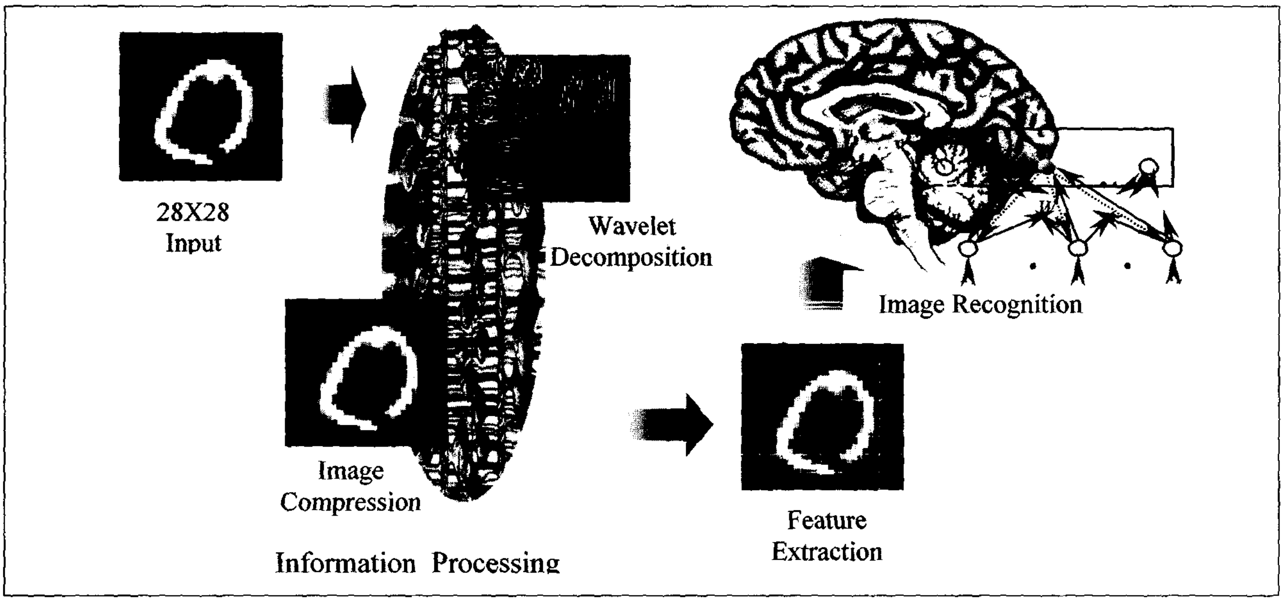


Fig. 3 Proposed Information Processing Model

To be able to explain certain objects, after reflecting on the relationship among them, we explain them in relation with other objects, For example, when we give a direction to our office to a third person, we describe the location of our office in relation with impressive buildings (big buildings or representable buildings) near the office. We then tell the third person how to get to the office by giving more concrete objects such as the room number or the room name. Like this, human is information processing about objects (what) with importance, whereas information about background (where) is processed less compared to the information about objects (what). Machine vision systems, however, do not make such distinction and process all the images. In recognizing objects, they exhaust the same amount of expenses on unnecessary background images as on objects, and the efficiency of storage space is low as well. In addition, if there are any changes in the background with no changes in the objects, unnecessary calculation expenses occur because new processing for the background is needed.

A mechanism, which processes input information from the retina, is found to be similar to the way that the wavelet transform processes input information. Information of low-

We implemented the image recognition model in figure 3.

B. Image Recognition Algorithms

Along with the development of research on artificial intelligence, image recognition studies have been successfully conducted by many areas, such as the areas of industry, military, and aerospace. Because neural networks have abilities of learning abilities and association memory, they are excellent in processing ambiguous signals that are changing continuously. Pattern recognition can be largely divided into two: 1) problem of finding specific information automatically from given data; 2) problem of understanding characteristics of given data by classifying them into more than two groups. The pattern recognition is being applied to various areas such as automatic interpretation of medical image information, letter and voice recognition, as well as biometrics of fingerprints, iris, and face recognition. To process human visual information, letter recognition is a necessary problem to solve.

Currently many algorithms have already been developed. Each algorithm has its own characteristic application areas. Generally, in case of classifying given data into clusters that have similar characteristics, the most commonly used methods are: Adaptive Resonance Theory (ART), Self-Organizing feature Map (SOM), and Fuzzy-ART. Among these methods, the algorithms of Grossberg's ART2 and Kohonen's SOM are chosen for this study. They are unsupervised learning algorithms that they learn on their own even if correct answers are not given to input patterns, which is similar to the biological recognition system.

Unsupervised learning algorithm is a fast performing model that can process learning in real time, has excellent characteristics of adaptation to various changes, and can be applied to both binary input and analog input patterns [7][10]. Connection-weighted changes of these unsupervised learning algorithms take an average value of input patterns and respond evenly to the generation of clusters; they are usually used for recognizers.

In this study, therefore, self-learning model is used as a recognizer and evaluates the performance of a model that implements a mechanism occurring in the retina.

IV. RESULTS OF EXPERIMENT

In this study, a proposed image model, which is based on the human visual information process, is implemented by using Visual C++ 6.0 in the environment of Pentium 1.7GHz, 256MB memory, and Window XP. We used MNIST database of the AT&T Corp. which it was used in a lot of paper on handwritten off-line number in figure 4 at this paper and tested. It is MNIST database newly compounding database of original NIST (National Institute of Standards and Technology). This is database using handwritten digit 0-9 in learning and recognition. NIST database is the binary image which was normalized with 20X20 sizes while keeping horizontal vertical ratio. The MNIST fitted size to 28X28 with the 8-bit gray-scale image that anti-aliasing handled database of NIST. The training data set selected 5,000 in 60,000. And, as for the test data set, a random sampling selected 5,000 in 10,000. And the person who wrote a number is not same.

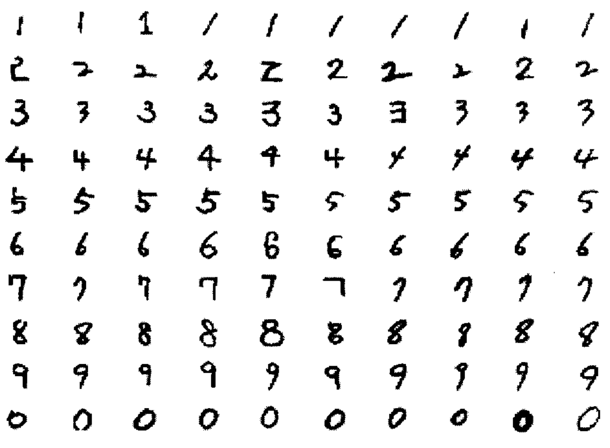


Fig. 4 MNIST Database Handwritten Digits

A 28×28 digit image was compressed with Daubechies (9,7) filter that is often used for the loss compression. For an experiment on the recognition process, the most commonly used neural network algorithms, ART2 and SOM, were used to suggest a recognition model. The results of the experiment are presented in Table 1.

Table 1 Experiment of off-line number recognition

bpp	MSE	PSNR	Recognition Rate
0.75	0.9040	48.5689	93.45%
0.50	0.9040	48.5689	91.66%
0.25	5.7965	40.4992	88.69%
0.125	19.0545	35.3308	88.09%
0.0625	49.1714	31.2137	85.11%

In the ART2 algorithm, 5,000 digits were recognized by the general recognition model with a recognition rate of 93.69%. The recognition rate of the proposed model was similar to that of a general recognition model with unprocessed input data. The results of ART2 or SOM showed almost the same recognition rate.

The memory of human beings fades away with time. Human recognition ability also decreases over time, and the accuracy of recognition becomes low. In addition, due to the limited capacity of the brain, a lot of images cannot be remembered, so human beings compress and recollect past memory. In existing machine vision models, for mistaken recognition and data management, input data were stored in DB. If we store all the data to DB and backup them, it will cost a lot of money. Therefore the size of images should be considered for maintenance. The size of input data is 28×28, and a necessary storage space to recognize 5,000 input data is 3,920,000 bytes. If this data are processed over time, a necessary storage space shown in Figure 5 will be increased exponentially.

In the experiment of this study, different compression rates were used over time by taking into consideration of human reflection functions. As it is shown in Table 1, recognition rates gradually decreases with compression rates, whereas the efficiency of storage space gets better, as shown in Figure 5. Therefore, in a recognition system considering human vision, the use of storage space is efficient as input data are processed according to the time flow.

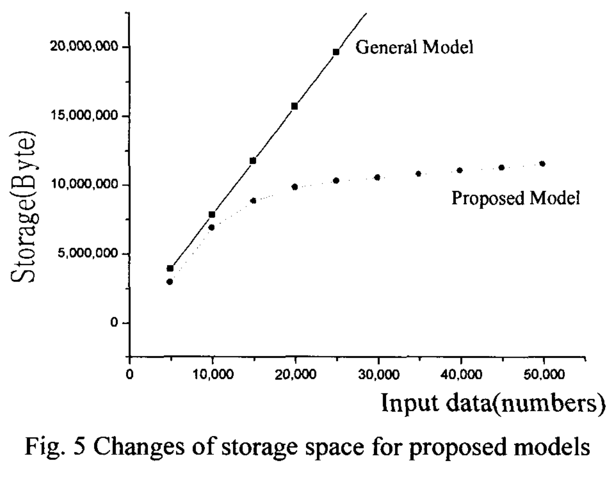


Fig. 5 Changes of storage space for proposed models

In the human visual information processing system, from the retina to ganglion cells, a substantial amount of information is processed by a mechanism. Human beings, however, do not feel any problems during the recognition process. When processing input data, minimum amount of important information (what) is lost, and maximum amount of less important information (where) is lost. This study implemented mechanism processing human visual information and examined its performance. The results showed that the recognition rate of the mechanism was not much different from that of the general recognition model, and the use of its storage space was found to be excellent.

V. CONCLUSTION

This study implemented mechanism information processing from the retina to ganglion cells, based on the human visual information processing. The results of experiment showed that there was not difference between a model considering a mechanism that processes visual information based on the human visual system and a general model in terms of their capability, when time was not taken into account. However, decreased efficiency of storage space for the general model over time must be considered. This study proposed an image recognition model considering the characteristics of the human visual system, not a strategic model applying an existing recognition model. As human reflection capability decreases over time, human beings need mechanism information processing to overcome the limited use of storage space. Therefore, after implementing this mechanism and comparing it with a general model, the two were found to be not different in terms of recognition rate, but the efficiency of storage space for the mechanism was found to be excellent.

If the modeling from ganglion cells to the primary visual cortex as well as the modeling from primary visual cortex to cognition process is completed in the future, real recognition performance will be further improved. The proposed model can be applied to the tracking of image recognition and to the manufacturing of artificial eyeball, and the potential of this model looks great.

REFERENCES

[1] M. fisxhler, O. Firschein, Intelligence: The eye, the brain and the computer, *Addison-Wesley*, 1987.

[2] R. Arnheim, Translation by Jung-oh Kim, Visual thinking, *Ewha Womans University Press*, 1982.

[3] J. Cho, "Research on data construction and recognition of artificial neural networks by fractal coefficients," *Dissertation at the Pusan National University*, 1998.

[4] Brain Science Research Center, "Research on Artificial Audiovisual System based on the Brain Information Processing," Research Paper by The Korea Advanced Institute of Science and Technology, *Department of Science Technology*, 2001.

[5] I. Lee, "Eyes and Computer," *Computer World*, 1989.

[6] I. L, "Human Beings and Computer," *Ggachi Gulbang*, 1992.

[7] D. Kim, "Neural Networks: Theory and Application (1)," *Hi-Tech Information*, 1992.

[8] J. M. Shapiro. "Embedded Image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing*, Vol. 41, No. 12, pp. 3445-3462, Dec. 1993.

[9] S. Mallat, "Multi-Frequency Channel Decomposition of Images Wavelets Models," *IEEE Trans. on Information Theory*, Vol. 11, No. 7, July 1992.

[10] S. Haykin, "Neural Networks: A Comprehensive Foundation," *MacMillan*, 1994.

[11] S. Hong. "JPE2000 Foundation for Compressed Coding of Still images," IDEC2002 Lecture, *IDEC*, 2002.

[12] R. C. Gonzalez, R. E. Woods, Digital image processing, Second edition, *Prentice Hall*, 2001.

[13] Y. LeCun, L. D. Jackel, L. Bottou, C. Cortes, J. S. Denker, "Learning algorithms for Classification: A Comparison On Handwritten Digit Recognition," in Neural networks: *The Statistical Mechanics perspective*, 1995.

[14] W. H. Dobelle. "Artificial Vision for the Blind by Connecting a Television Camera to the Visual Cortex," *ASAIO journal*, pp. 3-9, 2000.

[15] C. S. Burrus, R. A. Gopinath and H. Guo, Introduction to Wavelets and Wavelet Transforms, *prentice hall*, 1998.

[16] S. Mallat, A Wavelet tour of Signal Processing, *Academic Press*, 1998.

[17] J. M. Zurada, Introduction To Artificial Neural Systems, Boston: *PWS Publishing Company*, 1992.

[18] Darpa, Neural Network Study, AFCEA International Press, 1988.

[19] M. L. Minsky, and S. A. Papert, Perceptrons, Cambridge, MA: *The MIT Press*, First edition, 1969, Expanded edition, 1988.

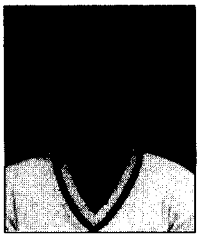
[20] T. Kohonen, Self-Organizing Maps, Berlin: *Springer-Verlag*. First edition was 1995, second edition 1997.

[21] K. I. Diamantaras, and S. Y. Kung, Principal Component Neural Networks, Theory and Applications, *NY: Wiley*, 1996.

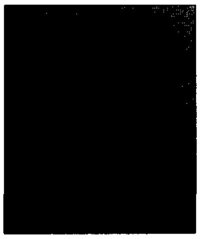
[22] S. Shah, M. D. Levine, "Information Processing in Primate Retinal Cone Pathways: A Model," TR-CIM-93-18, Centre for Intelligent Machines, *McGill University*, Montreal, December, 1993.

[23] S. Shah, M. D. Levine, "Information Processing in Primate Retina: Experiments and results," TR-CIM-93-19, Centre for Intelligent Machines, *McGill University*, Montreal, December, 1993.

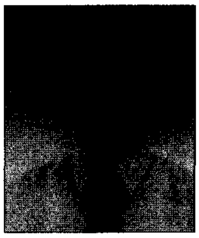
[24] J.E. Dowling. The Retina: An Approachable Part of the Brain. *Belknap Press of Harvard University Press*, Cambridge, MA, 1987.



Sung-Kwan Je
Received the B.S degree in computer science from Daegu University, Gyeongbuk, Korea, in 2000, and the M.S. degree and doctoral course in computer science from Pusan National University, Busan, Korea, in 2002, respectively. His research interests include Neural Networks and Application, Image Processing, Biological Signal Processing.



Jae-Hyun Cho
Received the B. S. degrees in Department of Computer Science from Pusan National University, Busan, Korea, in 1986 and the M. S. degrees in Department of Computer Science from Soongsil University, Seoul, Korea, in 1989, and the Ph.D. degrees in Department of Computer Science from Pusan National University, Busan, Korea, in 1998, respectively. From 1997 to present, he is an Associate Professor, Department of Computer Informaion, and Catholic University of Pusan, Korea. His research interests include Neural Networks and Application, Image Processing, Artificial Vision System.



Kwang-Baek Kim
He received the M. S. and the Ph.D. degrees in Department of Computer Science from Pusan National University, Pusan, Korea, in 1993 and 1999, respectively. From 1997 to present, he is an Associate Professor, Department of Computer Engineering, and Silla University in Korea. His research interests include Fuzzy Neural Networks and Application, Image Processing, Biological Signal Processing and Biomedical System.