Information Reduction in Retinal Region

Sung-kwan Je[1] , Kwang-baek Kim[2] and Jae-hyun Cho[3]

[1]Dept. of Computer Science, Pusan National University
[2]Dept. of Computer Engineering, Silla University
[3]Dept. of Computer Engineering, Catholic University of Pusan
Busan, Korea. E-mail: jimmy374@ pusan.ac.kr

***Abstract*— 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. 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.**

Index Terms— **Reduction, Retina, Human visual system, Information processing**

I. INTRODUCTION

Many of the properties of the human visual system are ultimately limited by the fidelity of image sampling within the retina and the strategies used to process the sampled information before its transmission to the cortex. The same may be said of machine vision systems. They are ultimately limited by properties of the sensor. Even at the retinal level, the performance of the human visual system far exceeds the capabilities of commercially available visual sensors and machine vision systems, and thereby motivates the study of human or primate retinas [22, 23].

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.

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 REDUCTION 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.

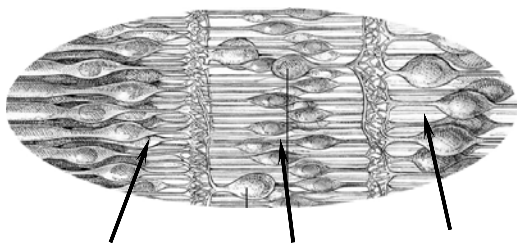

Fig. 1 Retinal Cells [22, 23]

A schematic diagram of the retinal cell structure is shown in figure1. The style of processing performed by the retina is dictated by two main requirements: (i) a large dynamic operating range and (ii) a reduction of data to meet the transmission bottleneck at the optic nerve. By utilizing two different photoreceptor types, (rods and cones), with differing sensitivities and adaptation mechanisms to extend their operating range, the retina is able to function over a very large dynamic range of illumination intensities.

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.

Fig. 2 General Recognition Processing of Computer Vision

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. With the continual expansion of multimedia and Internet applications, the needs and requirements of the technologies used, grew and evolved. In March 1997 a new call for contributions were launched for the development of a new standard for the compression of still images, the JPEG2000. This project was intended to create a new image coding system for different types of still images (bi-level, grey-level, color, multi-component), with different characteristics (natural images, scientific, medical, remote sensing, text, rendered graphics, etc) allowing different imaging models (client/server, real-time transmission, image library archival, limited buffer and bandwidth resources, etc) preferably within a unified system. This coding system should provide low bit-rate operation with rate-distortion and subjective image quality performance superior to existing standards, without sacrificing performance at other points in the rate-distortion spectrum, incorporating at the same time many contemporary features. The standard is intended to compliment and not to replace the current JPEG standards.

The wavelet transform has been adopted as JPEG2000 and has the following characteristics [11].

a) Superior low bit-rate performance
b) Continuous-tone and bi-level compression
c) Lossless and lossy compression
d) Progressive transmission by pixel accuracy and resolution
e) Random codestream access and procession
f) Robustness to bit-errors
g) Open architecture
h) Sequential build-up capability (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 Reduction in Retinal Region*

Although commercially available machine vision sensors are beginning to approach the photoreceptor densities found in primate retinas, they are still outperformed by biological systems in terms of dynamic range, and strategies of information processing employed at the sensor level. This has sparked recent interest in studying biological retinas in order to learn more about their information processing strategies with the hope of using this knowledge to design better machine vision sensors.

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.

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.

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-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. We implemented the image recognition model in figure 3.

$$c_k^{J-1} = \sum_{m=0}^{N-1} h(m) c_{M+2k}^{J} \qquad (1)$$

$$d_k^{J-1} = \sum_{m=0}^{N-1} g(m) d_{M+2k}^{J} \qquad (2)$$

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.

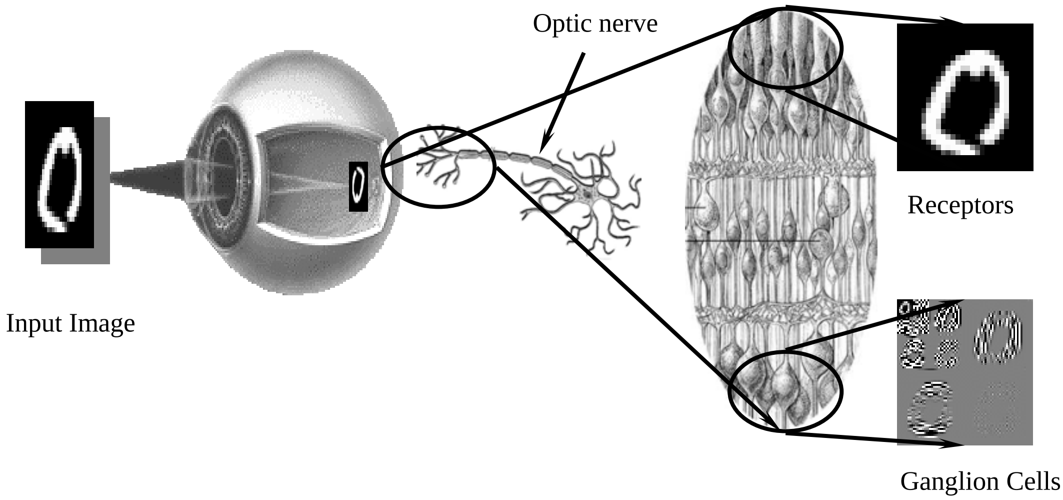

Fig. 3 Proposed Information Processing Model

If a discrete signal is $c_0^M, c_1^M, \cdots, c_{N-1}^M$, $N = 2^M$ from sampled to original signal f(t), we used equation (1) or (2) to show that wavelet transformed original image and calculated low-frequency, high-frequency filter of wavelet [11]. We applied low-frequency filter and high-frequency to a vertical direction and shared band with horizontal direction on original image and applied twice of filtering repeatedly again on a low-band section and use 3 level mallat tree to divide high-band section. Because 98% of high-band does division, we considered a diagonal, vertical, horizontal direction and did zigzag scan and did encoding in order to do encoding. We used the image that we passed through a decoding process and a reverse-quantization process again, and it was compressed encoding image on an image recognition process.

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 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.

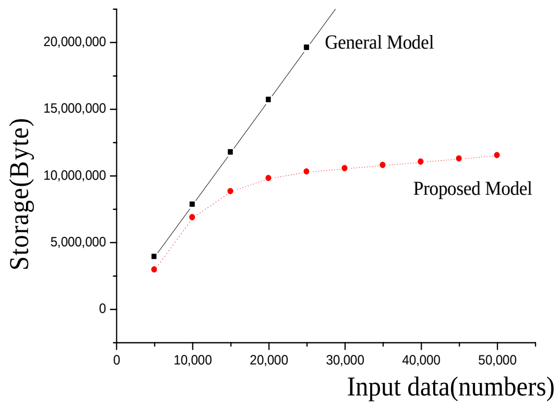

Fig. 5 Changes of storage space for proposed models

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.

[25] A. N. Skodras, C. A. Christropoulos and T. Ebrahimi, "JPEG2000: The Upcoming Still Image Compression Standard, " *Conference on Pattern Recognition*, pp.359-366, May 2000.