Digital Modelling of Visual Perception in Architectural Environment

Seo, Dong-Yeon* Lee, Kyung-Hoi**

Abstract

To be the design method supporting aesthetic ability of human, CAAD system should essentially recognize architectural form in the same way of human. In this study, visual perception process of human was analyzed to search proper computational method performing similar step of perception of it. Through the analysis of visual perception, vision was separated to low-level vision and high-level vision. Edge detection and neural network were selected to model after low-level vision and high-level vision. The 24 images of building, tree and landscape were processed by edge detection and trained by neural network. And 24 new images were used to test trained network. The test shows that trained network gives right perception result toward each images with low error rate. This study is on the meaning of artificial intelligence in design process rather than on the design automation strategy through artificial intelligence.

Keywords : Visual Perception, Digital Image Processing, Architectural Form, CAAD

1. Introduction

The Interface between human and objects can be observed by three dimensions as physical compatibility, cognitive compatibility, aesthetic compatibility. A large number of studies concerned with physical compatibility have been made by various numerical methods. Physical compatibility is mixed with cognitive compatibility and changed into image(aesthetic compatibility) of observer, but the aesthetic compatibility can not be able to evaluated only by numerical calculation1). For the reasons mentioned above, the architectural evaluation concerned with aesthetic dimension has mainly depended on the statistical methods2) analyzing verbal expression of subjects. However, at the point of visual perception, it its not easy for the evaluation merely by statistical methods to handle newly appearing variables from the ever-changing visual environment.

The evaluation of view, in addition, which is quite dependant on visual aesthetic perception3) may make a wrong estimate of human response when the elements considered for evaluation are restricted to physical aspect except for aesthetic aspect. It is a good example for such wrong estimate that residents in upper stories show no higher preference rate on view than in lower stories in high-rise apartment(upper stories: 33.4%, lower stories: 23.4%), though the conditions considered for numerical evaluation methods of upper stories are better than that of lower stories4).

Therefore, multivariate analysis including not only physical dimension but also aesthetic dimension and harmonious design reasoning methods are needed for proper evaluation of the objects depending on visual and aesthetic concept. Futhermore, computer can

* Graduate course of architectural. eng. in Yonsei Univ.
** Prof. of architectural. eng. in Yonsei Univ.

1) 中森 義輝, 感性 データ 解析, 森北出版株式會社, 2000, pp. 3~19.
2) Statistical methods mentioned is defined in this study as the first or second method type of aesthetic engineering; the first type is to build the foundation of translation date between feeling and expression, the second type includes the relation with subjects' life style and the third type is to make numerical model regarding subjects' expression as target function.
3) Arthur, L. M., T. C. Daniel, & R. Boster, “Science Assesment”, Landscape Planning 4, 1977, pp. 109~129
4) Park, Chan-Kyu, “A Study on the Design of Dwellings-on-Ground in the Flat Housing Estate”, Journal of the Architectural Institute of Korea, 1994. 4.

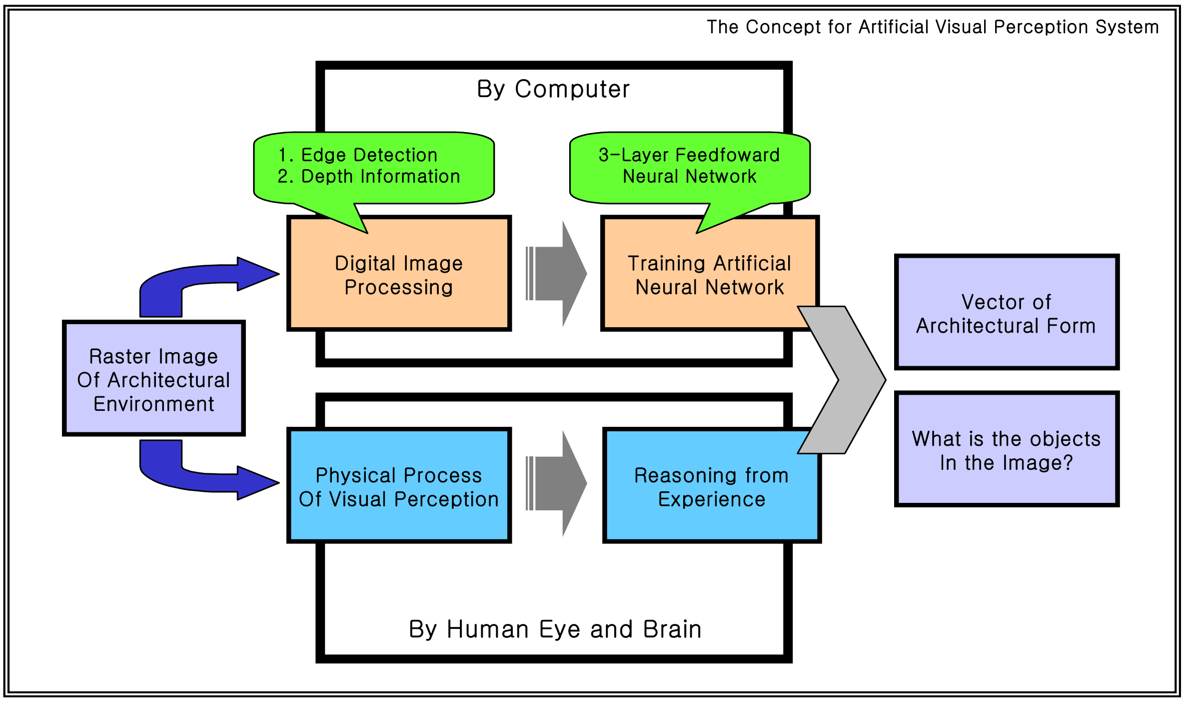


Figure 5. Comparative Concept for System Architecture

recognize the elements in pictures by similar process with the perception process of human firstly to construct computational system performing every step of aesthetic evaluation[5].

Current CAAD system can be used only by users' own effort of input and handling data and the forms of expression is too limited to support emerging figures and visual patterns[6]. With these issues in mind, this study intended to investigate the primary processing method to take the new turn to the third type of aesthetic engineering and to give progressive freedom onto CAAD environment.

Accordingly, this study begins with the hypothesis that it is most pressing problem to establish CAAD system supporting emerging patterns and aesthetic side that the system should have the function recognizing the characteristics of architectural environment. And the hypothesis is based on the visual perception theory that feeling toward visual objects can be explained and depicted if the characteristics in objects are discovered[7].

It is not so simple to recognize architectural form as letter recognition which can be done by training with some samples of simple property[8]. Visual perception structure of human should be understood and reflected in the system to recognizing architectural form.

As the first step in this study, theoretical study on the way how visual cues of architectural environment are changed to recognition, the structure of visual perception separated into two parts of low-level vision and high-level vision to find similar digital methods with each part.

Sobel Method of edge detection method is selected as the reconstruction strategy for low-level vision and Feed-forward Backpropagation Neural-Network is chosen for high-level vision. Various images of buildings, trees and landscape are digital processed and used for training network, the network is tested by sample images.

The purpose of this study is to show the skeleton method to recognize visual architectural environment as fundamental method making progress to aesthetic evaluation of visual environment. And this study intends to show the possibility for AI to be a useful design method[9] in architecture.

2. Recognition of Architectural Environment

5) Nagamachi, M., Image Technology Based on Knowledge Engineering and Its Application to Design Consultation, Ergonomics International, 1988.
6) Gero, J. and Yan, M., "Shape Emergence by Symbolic Reasoning", Environment and Planning B: Planing and Design, V.21, pp. 191~212, 1994.
7) R. Arnheim, Art and Visual Perception, University of California Press, 1974.
8) Rejean Plamondon, Heng Da Cheng, Pattern Recognition, World Scientific, 1991, pp. 115~138
9) E. Gombrich, Pictorial instructions, in: H.Barlow, C. Blakemore, M.Weston-Smith, Images and Understanding, Cambidge Univ. Press, Cambridge, 1990.

The massages from environment are changed into images through the process of human brain. Architectural environment is recognized by the images which is defined as the mental feature made from external experience[10].

The images from architectural environment play the role of primary translation for symbolic representation, understanding expression among objects and grasping meaning. The symbolic meaning of architectural environment is transferred through the physical form and to realize architectural environment is to generate image effectively[11].

According to Arnheim, visual image by which recognition of architectural environment is made is complete reconstruction of objects.

Generation of visual image needs visual cues by which people think and act.

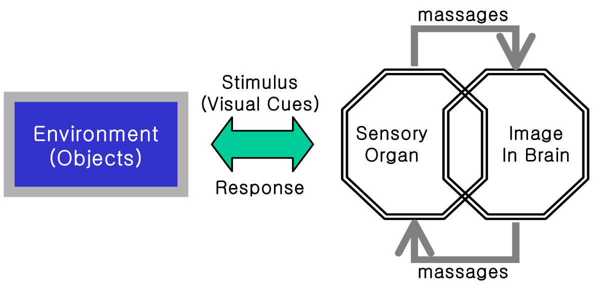


Figure 6. Recognition Process of Architectural Environment

Visual cues from external appearance of buildings consists of such things as entire feature, size, arrangement of windows and horizontal · vertical orientation of building[12]. The preference to environment can be estimated if architectural characteristics and visual cues are found(Kelly, 1955).

3. Digital Methods for Visual Perception

Visual perception is divided into two steps as the first term step and the latter term step, the first term step(low-level vision) works extracting forms from visual scene and the latter term step(high-level vision) does recognizing forms[13]. Each level of vision is reconstructed by computer programming using C++. This study gropes for adaptive methods to each step of visual perception.

10) Kenneth Bouling, The Image, Michigan Press, 1982.
11) Charles Burnette, The Mental Image and Design, Design for Human Behavior, Mcgraw Hill, 1977, pp. 169～172.
12) Martin Krampen, The Recognition of Building Function, EDRA 8, 1976, p. 404.
13) John Robert Anderson, Cognitive Psychology and Its Application, W. H. Freeman & Company, 1995, pp. 46～67.

3.1. Low-level Vision

The role of low-level vision is to change image or pattern shown by intensity of light to metaphysical expression.

Local information of light reflected from objects is gathered by receptor in the first step of low-level vision. The next step, then, is looking for descriptor from local light intensity referring to intensity information of neighbor pixels.

Lettvin has studied what makes the information of moving objects go to brain in the experiment of frogs, and Kuffler has investigated the ability which let cats and apes to recognize local contrast. Much ink has been spent on the studies showing that local contrast forms contour information which confirms the area of image along with depth information and structure information together. According to Arnheim the Gestalt psychologist, human apt to simplify objects to recognize immediately. In other words, the stimulus of external environment are to be seen as simple as possible according to the rule of visual perception. Thus the contour of forms in images which can be extracted by edge detection method is important and primary information containing visual cues to recognize objects.

Edge detection method is able to be satisfying method to reconstruct low-level vision consequently.

3.2. High-level Vision

As can be seen in Gestalt theories, reality has its own characteristics that is recognized certainly. Seeing external world is the interaction between observing subjects and the attribute of objects.

The role of high-level vision is analyzing objects by the information generated from low-level vision. Establishing visual perception system needs what it is rather than what looks like.

Sometimes low-level vision gives wrong information as incomplete or meaningless or distorted information.

A feature in image might be shade or a part lying on another or something out of knowledge base. Visual perception method catching hold of the meaning by context which is came from background and experience is that of human. Korean thatched houses and tile-roofed houses can not go with the scene of german province for example, condition and background are needed to recognize contextually.

To know expectation can play a crucial role to recognize environment when the image is too confused to be understood easily.

Therefore recognizing the reality of architectural form comes to be regarded as expectation of the image, which is similar concept with target value used for training neural network. Training neural network is the process of activating neurons by target value in brief, which models after the information process of human brain.

Neural network is able to be satisfying method to reconstruct high-level vision consequently.

3.3 Edge Detection Method as Low-level Vision

Contour of objects is important element for vision recognition and edge detection is crucial part of computer vision[14].

Contour tracking takes tow steps, the pixels creating contour are found in the first step and in the second step lines are generated with the pixels found. The pixels creating contour can be named edge. Tracking contour can be done with various operators given below[15].

(1) approximation of contrast value by the first or the second derivative differentials(Roberts 1965, MacLeod 1970, Marr and Hildreth 1980, Canny 1983)

(2) calculation edge model parameter(Hueckel 1971, Haralick 1980)

Edge is generally sought by the direction or the amount of contrast gradient

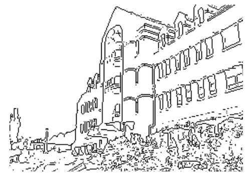
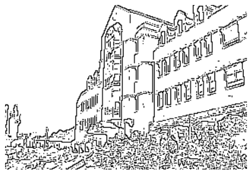
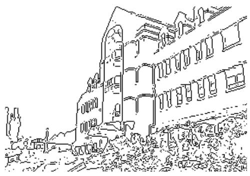
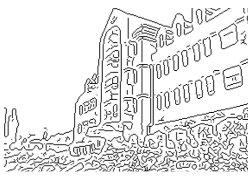
There are various methods from conventional method like contour following to dynamic programming for tracking contour. Contour tracking produces the set satisfying given conditions or the formula fit to lines.

In special case, Hough transformation, which is geometrical method to track the contour of circle and ellipse, can be turned to account[16].

Many studies have been made on edge detection methods, that is gradient methods, surface fitting methods and second-derivation methods etc. Sobel method and Canny method, that belongs to gradient methods, are prevalent methods among many edge detection methods.

Sobel method has the advantage to contain histogram equalization process and Canny method takes two threshold so that it is strong to noise[17].

Table 2. Comparison among Edge Detection Methods

Original Image	
Prewitt Method	
Roberts Method	
Sobel Method	
Canny Method (no threshold)	

Sobel method is selected in this study.

Sobel method is one of gradient methods, pixel composing contour is selected when the contrast value of the pixel exceed threshold.

To get gradient value from pixels,

$$Gf(x,y) = i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} \qquad \text{Eq. 1}$$
$$i = unit\ vector\ of\ x\ direction$$
$$j = unit\ vector\ of\ y\ direction$$

gradient mask extracted from Eq. 1 is used. For example, assuming that contrast value is given below in 3×3 pixel image,

f(x-1,y-1)	f(x,y-1)	f(x+1,y-1)
f(x-1,y)	f(x,y)	f(x+1,y)
f(x-1,y+1)	f(x,y+1)	f(x+1,y+1)

the gradient value(Gx) is gained from Eq. 2,

$$G_x = f(x+1,y+1) + 2f(x+1,y) + f(x+1,y+1) - f(x-1,y-1) - 2f(x-1,y) - f(x-1,y+1) \qquad \text{Eq. 2}$$

and Gy is gained from Eq 3.

14) Attneave F., Some Informational Aspects of Visual Perception, Psychological Review 61, 1954, 183~193.

15) L. S. Davis, “A Survey of Edge Detection Techniques”, Computer Graphics and Image Processing, 1975, pp. 248~270.

16) Parker, James R. Algorithms for Image Processing and Computer Vision. New York: John Wiley & Sons, Inc., 1997. pp. 23-29.

17) Canny, John. "A Computational Approach to Edge Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986. Vol. PAMI-8, No. 6, pp. 679-698.

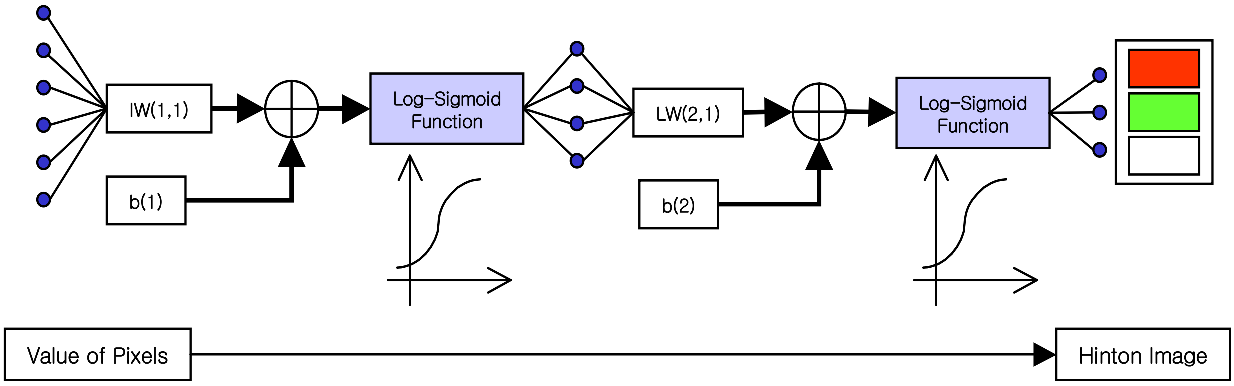


Figure 12. Architecture of Neural Network System

$$G_y = f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1) - f(x-1, y-1) - 2f(x, y-1) - f(x+1, y-1) \quad \text{Eq. 3}$$

Gradient value is G,

$$G = \sqrt{G_x^{\,2} + G_y^{\,2}} \quad \text{Eq. 4}$$

Direction of contour line at pixel(x,y) is θ.

$$\theta = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad \text{Eq. 5}$$

When G exceeds T(threshold value),

$$G > T$$

pixel(x,y) can be regarded as the point on the contour line.

3.4 Architecture of Neural Network

The study of Kuffler(1953) shows how information is encoded by cells of neurons. Cells are activated for themselves with invariable rate and some cells of neurons are activated vigorously when light stimulate the retina.

Cells show reaction at their maximum if contrast make lines according to Hubel & Wiesel(1962).

Neural network model is that nodes connected with neurons are represented as directed graphs with arrows. Each node has value which mean connecting strength of synapse in brain. There are two kinds of nodes such as input nodes and output nodes in neural network, input nodes send the stimulus to network and output nodes represents calculated results.

Learning by neural network is done by methods controlling weights and is the process of repeated training which calculate input function. The weight of nodes is not proper before training. Regulation of weight stops when the train is finished.

Many kinds of neural network have been devised and can be classified as table 3 due to pattern of input and output and the way of learning.

For pattern recognition, Neural network has the advantage that it can be used in case of impossibility with statistical methods. In other words, the information of probability does not have to be made to recognize forms[18].

Multi Layer Feedfoward Back-Propagation Neural Network is selected for this study, which acts on multi form recognition better than any other methods, consists of decisive functions and layers.

Table 4. Classification of Neural Network

Neural Net. Classifier	Binary Input	Supervised	Hopfield Network	
			Hamming Network	Optimized Classifier
		Unsupervised	Carpenter/ Grossberg Classifier	Clustering Algorithm
	Consecutive Value	Supervised	Perceptrons	Gaussian Classifier
			Multi Layer Perceptrons	K-nearest Neighbor
		Unsupervised	Kohonen Self Organization	K-means Clustering

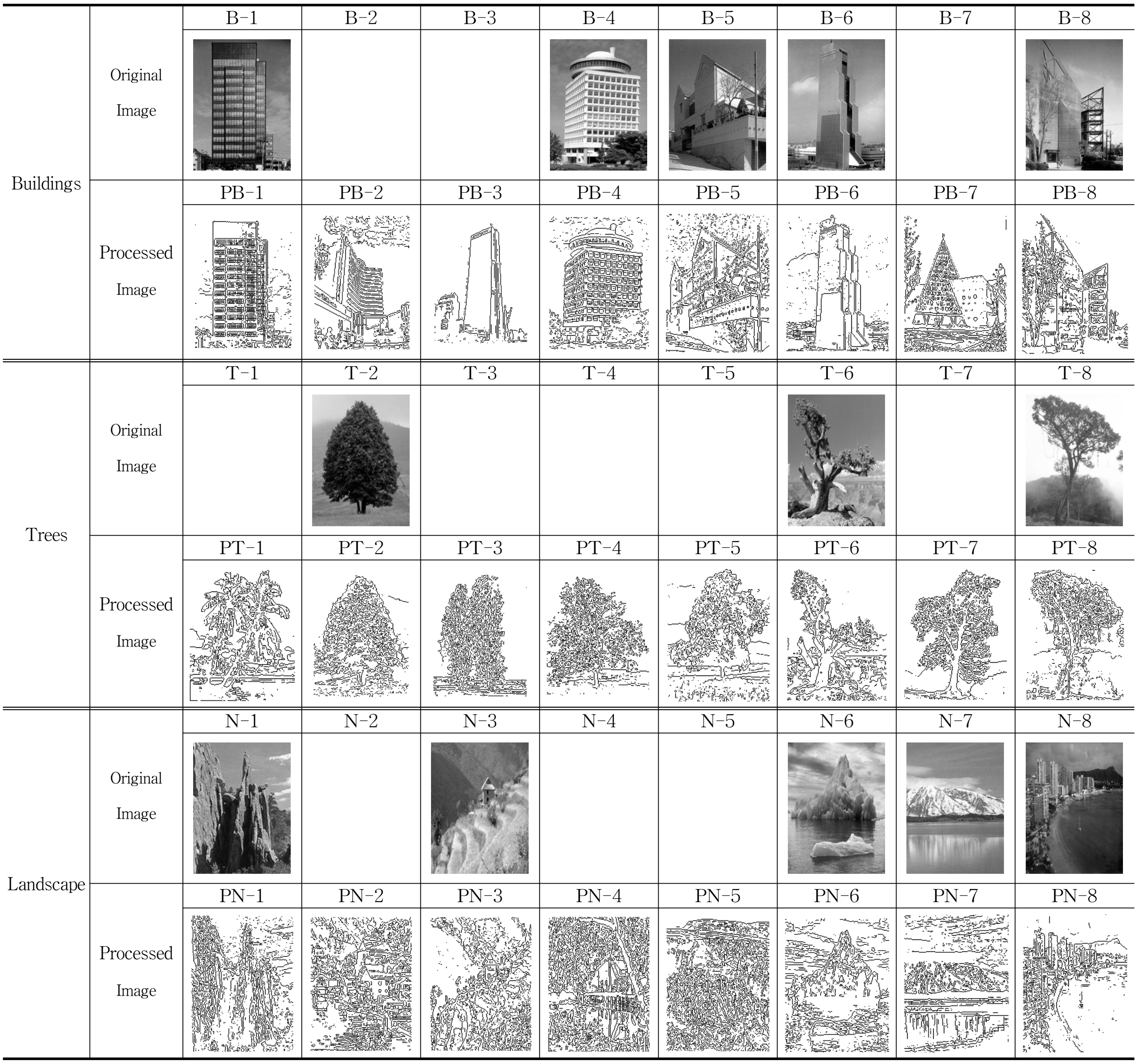
Figure 3 shows the network structure fo this study.

The number of input neuron is 30000 due to the size of images(150×200 pixel) and output display is designed to 3 neurons. The number of mid-layer is 36, which is selected by experiments. It takes long time to construct network if mid-layer has many neurons. It is hard to get convergence of error if mid-layer has a few neurons. Log-Sigmoid function is selected as activation function, Resilient Backpropagation as training methods.

Resilient Backpropagation is effective method to control computer memory because controlling memory should be given much more weight than processing

18) Rafaedl C.Gonzalez & Richard E. Woods, Digital Image Processing, Addison Wesley Longman, 1992.

Table 5. Images for Training



speed. Iteration of training goes to 19 times until error meets 10^{-6}.

The network for this study is designed to give three values as the result of recognition.

The result of recognition is represented in Hinton image and value then.

Hinton image is the displaying method which have the results understood at a glance. The size and colour in Hinton image are varied according to the activation rate of neurons.

4. Learning and Testing Forms

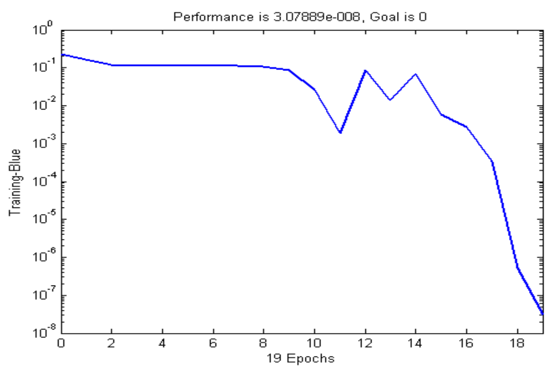


Figure 61. Falling Error

Every eight images of buildings, trees and landscape[19] are processed by edge detection method. Processed

Table 6. Test Result of Trained Network

Buildings		B-9	B-10	B-11	B-12	B-13	B-14	B-15	B-16
	Input Image								
	Result	[9.298×10^{-96}; 6.81×10^{-8}; 0.00024433]	[8.3879×10^{-122}; 8.6948×10^{-6}; 3.1143e-011]	[9.5947×10^{-96}; 6.7866×10^{-8}; 0.00024511]	[9.0019×10^{-96}; 6.8228×10^{-8}; 0.00024298]	[2.9661×10^{-107}; 5.237×10^{-7}; 5.7834×10^{-8}]	[9.0242×10^{-96}; 6.8237×10^{-8}; 0.00024302]	[1.4933×10^{-105}; 1.8938×10^{-5}; 2.2971×10^{-8}]	[8.1117×10^{-92}; 3.5428×10^{-6}; 7.6492×10^{-5}]
Trees		T-9	T-10	T-11	T-12	T-13	T-14	T-15	T-16
	Input Image								
	Result	[7.4183×10^{-117}; 2.0699×10^{-25}; 1.0000]	[6.8369×10^{-145}; 3.5245×10^{-50}; 1.0000]	[3.6393×10^{-139}; 2.512×10^{-45}; 1.0000]	[6.53×10^{-136}; 8.1138×10^{-43}; 1.0000]	[6.6373×10^{-145}; 3.379×10^{-50}; 1.0000]	[3.0924×10^{-144}; 1.2454×10^{-49}; 1.0000]	[6.0035×10^{-116}; 6.5564×10^{-25}; 1.0000]	[1.5538×10^{-111}; 7.0032×10^{-20}; 0.99961]
Landscape		N-9	N-10	N-11	N-12	N-13	N-14	N-15	N-16
	Input Image								
	Result	[2.649×10^{-147}; 0.99949; 1.0922×10^{-22}]	[2.649×10^{-147}; 0.99949; 1.0922×10^{-22}]	[2.649×10^{-147}; 0.99949; 1.0922×10^{-22}]	[2.6491×10^{-147}; 0.99949; 1.0922×10^{-22}]	[2.6491×10^{-147}; 0.99949; 1.0922×10^{-22}]	[2.649×10^{-147}; 0.99949; 1.0922×10^{-22}]	[2.649×10^{-147}; 0.99949; 1.0922×10^{-22}]	[2.6491×10^{-147}; 0.99949; 1.0922×10^{-22}]

images by edge detection method are changed to binary array to train network.

Training error reaches to 3.07889×10^{-8} in 19th iteration.

Following table shows target value.

Table 7. Target Value

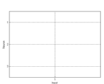
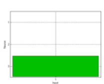
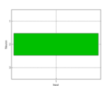
	Buildings	Trees	Landscape
Hinton Image			
Target Value	tv① = $\begin{bmatrix}0\\0\\0\end{bmatrix}$	tv② = $\begin{bmatrix}0\\0\\1\end{bmatrix}$	tv③ = $\begin{bmatrix}0\\1\\0\end{bmatrix}$

Table 7 shows each error of output neurons and also that images of trees are trained better than buildings and landscape.

19) landscape is defined as the concept containing natural elements of trees and artificial elements of buildings, and landscape is expected to generate different result from buildings and trees in this study.

Table 8. Error of Output Neurons(shown in 10^{-3})

B	B-1	B-2	B-3	B-4	B-5	B-6	B-7	B-8
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0025	0.0006	0.0129	0.0006	0.0002	0.0030	0.0001	0.0001
	0.1025	0.0012	0.0003	0.0012	0.0021	0.0009	0.2400	0.2418
T	T-1	T-2	T-3	T-4	T-5	T-6	T-7	T-8
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
L	N-1	N-2	N-3	N-4	N-5	N-6	N-7	N-8
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.5102	0.5131	0.5102	0.5102	0.5102	0.5102	0.5102	0.5146
	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

The training of trees' form does not give any error because trees have the standard of stem, branch and leaves. Forms of trees are varied within the range of the standard. Buildings have common feature of straight lines, but it is hard to find common features in landscape images comparatively. The results of N-13 and N-14 are remarkable. N-13 and N-14 are classified to landscape though they have some buildings.

5. Discussion

Images tested are different from each other, geometrical type and perspective aspect and so on are different. However, images can be classified properly due to common visual cues, buildings are recognized as what it is due to contour lines, windows and horizontal, vertical and parallel lines. The result of classification means that computer recognize visual cues mentioned by Krampen. Target value given to train of N-2, N-4 and N-8, though the images contains visual cues seen in building images, make network recognize N-13 and N-14. This means that the network resembles human contextual recognition.

The simple structure of stem and volume causes easier recognition than buildings and landscape.

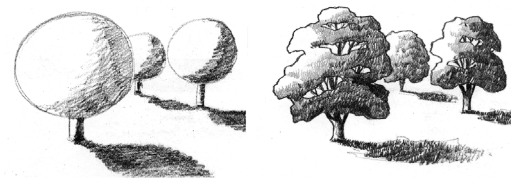


Figure 89. Concept and Derivation of Tree

Given that people are asked to only 'draw' buildings, trees and landscape, which subject they can draw easily? They could draw trees without complex thinking process because simple concept and derivation of tree is existed in brain in the form of trained neurons, which testifies that training neural network reconstructs high-level vision of human.

To be the design method supporting aesthetic ability of human, CAAD system should essentially recognize architectural form in the same way of human[20].

Fundamental method and possibility of making computer reason and learn from experience are investigated in this study.

Supporting the estimation of human sense and emotion, limitation of CAAD system could be removed to some degree when the function of human is reconstructed by proper ways.

References

1. 中森 義輝, 感性 データ 解析, 森北出版株式會社, 2000, pp. 3~19.
2. Arthur, L. M., T. C. Daniel, & R. Boster, “Science Assesment”, Landscape Planning 4, 1977, pp. 109~129.
3. Park, Chan-Kyu, “A Study on the Design of Dwellings-on-Ground in the Flat Housing Estate”, Journal of the Architectural Institute of Korea, 1994. 4.
4. Nagamachi, M., Image Technology Based on Knowledge Engineering and Its Application to Design Consultation, Ergonomics International, 1988.
5. Gero, J. and Yan, M., “Shape Emergence by Symbolic Reasoning”, Environment and Planning B: Planing and Design, V.21, 1994, pp. 191~212.
6. E. Gombrich, Pictorial instructions, in: H.Barlow, C. Blakemore, M. Weston-Smith, Images and Understanding, Cambidge Univ. Press, Cambridge, 1990.
7. Kenneth Bouling, The Image, Michigan Press, 1982.
8. Charles Burnette, The Mental Image and Design, Design for Human Behavior, Mcgraw Hill, 1977, pp. 169~172.
9. Martin Krampen, The Recognition of Building Function, EDRA 8, 1976, p. 404.
10. John Robert Anderson, Cognitive Psychology and Its Application, W. H. Freeman & Company, 1995, pp. 46~67.
11. Attneave F., Some Informational Aspects of Visual Perception, Psychological Review 61, 1954, 183~193.
12. L. S. Davis, “A Survey of Edge Detection Techniques”, Computer Graphics and Image Processing, 1975, pp. 248~270.
13. Parker, James R. Algorithms for Image Processing and Computer Vision. New York: John Wiley & Sons, Inc., 1997. pp. 23-29.
14. Canny, John. “A Computational Approach to Edge Detection”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986. Vol. PAMI-8, No. 6, pp. 679-698.
15. Rafaedl C.Gonzalez & Richard E. Woods, Digital Image Processing, Addison Wesley Longman, 1992.
16. Lee, Han-Seok, “An Analysis of the Visual Information Process in Architectural Design process”, Journal of the Architectural Institute of Korea, 1995. 9.

20) Lee, Han-Seok, “An Analysis of the Visual Information Process in Architectural Design process”, Journal of the Architectural Institute of Korea, 1995. 9.