19 A Symbolic/Neural Hybrid Approach to Emergent Subshape Recognition

Mon-chu Chen

Architecture Group, Institute of Applied Arts, National Chiao-Tung University, Hsinchu, Taiwan

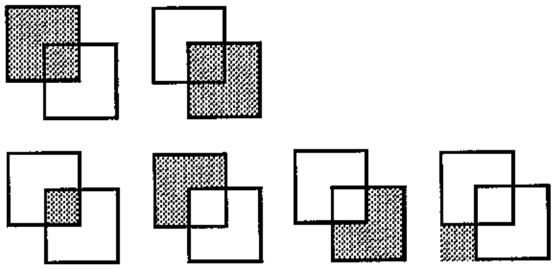
Recognizing emergent subshape is one kind of human visual behavior. People usually recognize several distinct emergent subshapes from primary shapes and give them different interpretations. This paper presents a symbolic/connectionist hybrid system to provide computers the ability of this kind. Through this approach, the recognition system is divided into three modules. Source images are sent to the first module, that is a connectionist network, of the hybrid system. The network is responsible for transforming the source image into abstract visual data, named Pre-attention Distribution and Local Feature Information. Then, the abstract visual data are processed in the second module that is a symbolic subsystem. The subsystem is responsible for making decision in the Visual Search Attention processes and for managing the features of the whole shape. Finally, another connectionist network takes the previous results from the symbolic subsystem and performs the final recognition.

INTRODUCTION

An emergent subshape means one kind of shape that is interpreted as different shapes other than the primary shape. These interpreted shapes in addition to the primary shape are called emergent subshapes. They are the shapes only implicitly/explicitly existing in an already interpreted shape. Recognizing emergent subshapes is one kind of human visual phenomenon. This phenomenon was investigated by the Gestalt psychologists in the very early of the 20th century [i]. It is very easy for people to see emergent subshapes in addition to the primary shapes in one source image [ii]. The final recognition of each person might vary. Since "SEEING" plays a very important role in design [iii], the fact of different recognition of a primary shape between people is used to explain the creativity of human design behavior. It plays an important role in the field of both creative design thinking and computer-aided design [iv]. People who deal with computer-aided design and would like to make computers more intelligent or human-like cannot avoid discussing this issue.

1.0 EMERGENT SUBSHAPES

One example of recognizing emergent subshapes in terms of emergent subshapes is shown in Figure 1: A picture formed by two overlapped squares. In ordinary CAD system, the picture can only be recognized as two squares. However, people could easily interpret it as not only the two original squares but also the smaller intersection square and L shapes, etc. A CAD system without this ability does not allow other possible reinterpretations. Consequently, users can only design under the situation of limited shape recognition. The issue also occurs in the field of Automatic Design. Most of the models of design thinking are based on Simon's search model [v]. Without the ability of emergent subshape recognition, this kind of rule-based system can only generate undoubtedly fewer design alternatives.



The researches on emergent subshape recognition have just been increased recently. Most of the researches are based on symbolism[vi-xi]. Although systems in this approach acquire good results, they lacks the abilities to process unclear or incomplete shapes. Recently, a connectionist network model was proposed [xii] to address this issue. Providing the power of tolerance, the model is still defective in both systematization and efficiency that belongs to major strengths of symbolic processing. Many researches on AI, and cognitive psychology support that integrating connectionist network and symbolism is the way it has to be [13][14]. Further more, some phenomena of seeing shapes point out that the recognition is a combination of low-level perception and high-level cognition. According to above viewpoints, a hybrid system approach should be a better solution.

2.0 SYMBOLIC APPROACH

Symbolic systems usually acquire good results on recognizing emergent subshapes. Sometimes, they did too well for people to handle. In Gero's example of symbolic approach [6], the emergent shapes are discovered by shape hiding and data-driven searching. During the shape hiding process, the original primary shape becomes a set of segments and vertices. After the shape hiding process, a data-driven search was performed to collect partial segments and vertices, which satisfied some constraints of specific shapes, of the set. Those constrains include topological, geometrical and dimensional considerations. Shapes fulfill these restrictions are defined to be emergent subshapes.

According to the algorithm Gero proposed, emergent subshapes can be quickly, efficiently and systematically recognized from primary shapes. But, the algorithm generates too many "emergent subshapes". Most of them seem less meaningful and hardly to be recognized from the original shape by people. Further more, the importance of every subshape candidate has no difference to each other. This is against the human visual behavior. Gestalt psychologist mentioned that different shapes in a picture will catch our attention with different degre[i]. People usually see some shapes easier than others. Gero's algorithm can hardly perform like that.

In addition to the above limitation, there is also, generally speaking, a defeat to every symbolic method. Since they are dealing with symbols, processing unclear or incomplete data become a very hard problem. Currently, shape representations in most CAD systems adopt symbolic methods. This helps symbolic approach of emergent subshape recognition to be adapted into these CAD systems, but it is still doubted whether or not future CAD systems deal with shapes symbolically. How could these symbolic CAD systems accept blurred, noised pictures someday? How could they deal with incomplete shape in the future? Will symbolic data structure be the only one kind of shape representation in CAD systems?

3.0 CONNECTIONIST NETWORK APPROACH

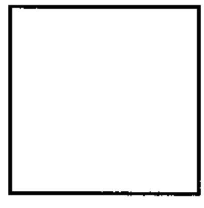
Connectionist networks, also called connectionist networks, have been addressed as a powerful tool in the field of artificial intelligent. The networks essentially attempt to model human brain functions by connecting artificial neurons in different ways. Thus, it is commonly believed that the developments of cognition and AI may go deeper from the connectionist point of view. Liu [12] first introduced this technique into the research of emergent subshape recognition. He took advantage of some results of the research on human visual behavior in Psychology as an analogy to his algorithm, such as Recurrent Attention and Searchlight Attention.

The idea of Recurrent Attention is that within a specific field of vision that includes multiple objects, people naturally and gradually focus on an object by paying more attention recurrently to the stronger stimulus. Because the recurrent attention idea still encounters some limitations, Liu proposes a searchlight attention mechanism to be embedded into the network. The source images are dealt systematically with searchlight attention to filter out the data outside the receptive field. The passed data are sent to the recurrent attention network for further processes.

Liu's model draws a significant result on the recognition of incomplete, unclear emergent subshapes. The simulation results include both explicit and implicit emergent subshapes. It also gives some conclusions that overcome the limitation of the previous symbolic model. But, there are some problems needed to be refined. The fact that the system could recognize a primary shape in several different interpretations is based on the usage of searchlight attention mechanism. It implies whenever the system focuses the same receptive field, there will be only one interpretation of this primary shape (Figure 2). Moreover, the searchlight attention mechanism is rather a systematic iteration. Usually people would not pay attention on a picture like a mechanical scanner.

Beside these considerations, there are two general limitations in every connectionist network systems:

- it needs lots of training processes
- once the network is trained, it will be very difficult to make the network to learn new shapes.



4.0 THE HYBRID MODEL

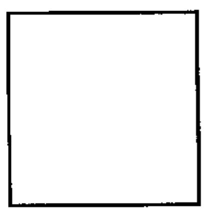
The hybrid system given here is mainly divided into three modules that are Pre-Attention Module, Visual Search Attention Module, and Recognition Module. The first module and last module are based on connectionist networks. The second one is a symbolic subsystem in association with a shape description knowledge-based system (KBS).

5.0 PRE-ATTENTION MODULE

The input data of the first model which is a connectionist network includes the source bitmap image, the current focusing coordinates, and the moving direction. And, the output data includes abstract visual data and local feature information. These outputs could be seen as a combination of the "Primal Sketch" in Marr's theory [xv] and the Dobson's "feature unit templates" [xvi]. The data stands for the stimulus distribution of the receptive field and the partial feature of the whole shape.

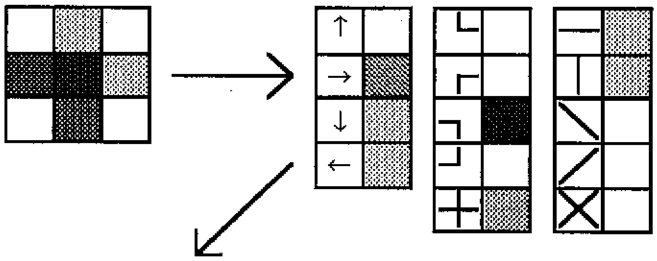
Regarding the abstract visual data, since the connectionist network was trained by the rules that the more times people pay attention on a portion of one picture, the higher activation that portion will be turned out. This abstract data of the processed image contains the stimulus strength of every pixel in the source image. The output bitmap is somehow like the image people maturely catch when they get a glimpse of the original picture (Figure 3). These stimulus distributions will be sent to the next module, that is the Visual Search Attention Module, to determine where the focusing receptive field will be moved to. Because the pre-attention involves the previous moving direction and focusing coordinates, it implies that the outputs might be different whenever the system focuses on the same receptive field.

The local features include types, orientations, and coordinates. The results in this module might be "a cross vertex at position (x1, y1)", or "a 45-degree segment at position (x2, y2)", etc. These features will be sent to the next module and stored in the short-term memory for feature processing. Since the network here is only responsible for the local feature recognition, the size of the network and the number of necessary training samples would not be as many as in Liu's model.



6.0 VISUAL SEARCH ATTENTION MODULE

This symbolic module consists of two parts: a saccade engine and a shape representation KBS. Saccade is a well-known human visual perceptual behavior when looking at pictures [xvii]. The saccade engine takes the output data from the first module and decides which portion of the original picture the following recognition process should focus on. At the mean time, the local feature information has also been stored in the short-term memory (Figure-4). Once the symbolic engine has focused on one spot, so-called the Fixation, the local feature information of the attended spots associated with the current coordinate will be compared to the pre-defined data in the shape representation KBS to then be sent to the third module of the recognition system.



Type	X	Y	Stimulus	
┼	1	8	10.33	
┐	0	0	6.20	

The saccade engine works with a global stimulus distribution map. The map records the stimulus weights of all portions which the recognition system has attended. Whenever the Pre-Attention Module outputs a local stimulus distribution, the engine adds them into the global map. Taking the global stimulus distribution map as an inverse weighted function, the engine multiplies local and global maps and finds the maxmun as the direction of the next movement.

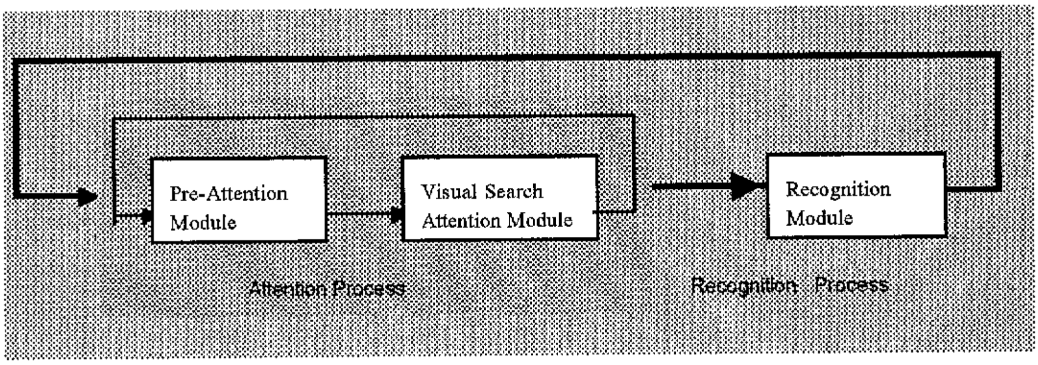
The knowledge-based system contains the shape topology constrains and shape representations, based on the local features. The KBS is responsible for inspecting the features of recently-visited spots stored in the short-term memory whether or not they match the constrains. If they are satisfied, the data will be sent to the next module, the Recognition Module, for the final process.

7.0 RECOGNITION MODULE

The last module is an ordinary connectionist network that can recognize different shapes and distinguish whether they are explicit emergent subshape, implicit emergent subshape, or complete shapes. It receives the visual codes, such as the number of vertices, related positions and angles, etc., then transfers them into verbal codes such as <rectangle>, <triangle>, etc. Since the network concerns with the features of the multiple shape rather than the shape itself, the network can recognize various sizes of the same shape.

8.0 ATTENTION PROCESS AND RECOGNITION PROCESS

During the recognition processes, the first two modules form a cycle which is named the attention process. In the cycle, the Pre-Attention Module continues generating local stimulus distributions and feature information, while the saccade engine of the Visual Search Attention Module directs the Pre-Attention Module to follow the path of attention search. The shape representation KBS also records the local feature information that the Pre-Attention Module generates. Whenever the saccade engine focuses on a vertex, the KBS begins to collect feature data in the short-term memory and sent these data to the third module. The attention process cycle embedded with the third module, that is the recognition process, forms another larger cycle. It keeps on focusing any possible recognizable portion of the source picture until it has finished scanning the whole picture or it is asked to stop (Figure 5).



DISCUSSION

Since this approach combines the symbolic method and connectionist network technique, it takes the advantages of both fields. Comparing to the Gero's model, this hybrid model could process unclear or incomplete shapes. The emergent subshapes generated by the hybrid model have different priority during the recognizing process. These two points are the major defeat of Gero's symbolic model.

Comparing to the Liu's connectionist network model, the Visual Search Attention Module is based on the symbolic approach, thus, the hybrid model overcomes some major weakness of the connectionist network research on emergent subshape recognition. In this hybrid model, the results would not necessarily be the same whenever the search attention focuses on the same portion of the whole picture. This is a problem that Liu's model has not solved. The path of attention search produced by the saccade engine is more complete than the Searchlight Attention in Liu's model.

The recognition method in this proposed hybrid model is feature-based. It provides many advantages, such as recognizing various size of the same shape and reducing large number of training samples. These are the major disadvantages of the Liu's connectionist network model.

Besides those considerations listed above, the hybrid system combines both symbolic and connectionist network methods to simulate human behavior. Thus, the recognition results are supposed to be closer to the human cognition behavior.

REFERENCE

[i] K. Koffka, *Principles of Gestalt Psychology,* New York: Harcourt, Brace and Co., 1935

[ii] Y. T. Liu, "Some Phenomena of seeing shapes in design", *Design Study,* v 16, n 3, pp. 367-385, 1994

[iii] D. A. Schön and G. Wiggins, "Kinds of seeing shapes and their functions in design", *Design Study,* v 14, n 3, pp. 135-156, 1992

[iv] W. J. Mitchell, *The Logic of Architecture,* MIT Press, Cambridge, MA., 1990

[v] H. A. Simon, *The sciences of the Artificial,* 2nd ed. MIT Press, Cambridge, MA., 1981

[vi] J. S. Gero, and M. Yan, "Discovering emergent shapes using a data-driven symbolic model", in U. Flemming and S. Van Wyk (eds), *CAAD Futures '93,* Elsevier, Amsterdam, pp. 3-17, 1993

[vii] T. Nagakura, "Shape recognition and transformation: A script-based approach", in M. McCullough, W. J. Mitchell and P. Purcell (eds), *The Electronic Design Studio,* MIT Press, Cambridge, MA., pp. 149-171, 1990

[viii] M. Tan, "Closing in on an open problem-reasons and a strategy to encode emergent subshapes, in J. P. Jordan (ed.), *ACADIA '90 Proceedings: From Research to Practice,* Montana State University, Big Sky, Montana, pp. 5-19, 1990

[ix] E. Edmonds and B. Soufi, "The Computational modeling of emergent shapes in design", in J. Gero and F. Sudweeks (eds), *Preprints of the Second International Round Table Conference on Computational Models of Creative Design,* 1992

[x] G. Stiny, "Weights", *Environment and Planning B: Planning and Design,* 19, pp. 413-430, 1992

[xi] M. Yan, "An Evolutionary Approach to Shape Emergence", in *Proceedings of the 1st IEEE Conference on Evolutionary Computation, Part 2,* IEEE, Piscataway, NJ., pp. 775-780, 1994

[xii] Y. T. Liu, "A neural-network approach to shape recognition and transformation", in U. Flemming and S. Van Wyk (eds.), *CAAD Futures '93,* Elsevier, Amsterdam, pp. 19-36, (1993)

[xiii] M. Minsky, "Logical versus analogical or symbolic versus connection or neat versus scruffy", *AI Magazine,* v 12, n 2, Summer, pp. 34-51

[xiv] Y. T. Liu, "Symbolic CAAD versus connectionist CAAD", *Language of Design,* n 2, pp 325-346, 1994

[xv] D. Marr, *Vision,* Maple-Vail Book, 1982

[xvi] V. G. Dobson, "Visual search with decrementing network competitive windows", in D. Brogan (ed), *Visual Search,* Taylor & Francis, London, 1990, pp 211-233

[xvii] R. H. S. Carpenter (ed), *Vision and Visual Dysfunction (8): Eye Movements,* CRC Press, Boca Raton, 1991