'92 KACC 1992. 10. 19~21

Personal Computer Networking for Concept Understanding

Y.Sakai, and M.Kitazawa

Dept. of Mechanical Engineering, Yamaguchi University, Ube, Yamaguchi 755, Japan

Abstract

This paper deals with the architecture of an information processing model for the human concept understanding required in constructing intelligent man-machine interfaces. The architecture employed is a parallel processing by networking. For this purpose, personal computers are interconnected by LAN and are, in their roles, divided into three levels. A concept has two aspects; i.e., language and image. In the present model, the system as the holistic whole of personal computers together with peripheral devices processes visual information in cognitive level, searching for feasible solutions from a variety of aspects. An image inputted through peripheral systems is categorized and matched with those ever experienced with the aid of that categorization, and thus an image is identified.

Introduction

A methodology for automatic concept understanding is necessary in order to construct an intelligent man-machine interface which can reduce human operator's load for communication. A concept has two aspects of language and image. Bilateral translation between these two sides. An this issue is essential in an intelligent interface. For this purpose here, imagerial aspect of concepts is dealt with extracting features of figures and letters from visual data. Thus obtaining the correspondence between images and words will cultivate the power for concept understanding in an interface. A computer networking is employed is employed in implementing a basic system for this purpose here. One of the major advantages of such an architecture is the parallel processing.

Description of the System and the Processing

1) The Present System

Mimicking the procedure in human visual information processing, a personal computer network system is formed as is shown in Fig.1. To each stage in cognition, sensory level to concept-understanding level, a device(s), a computer and/or peripheral systems, is assigned. More concrete descriptions for these tasks will be described below.

Local Area Network (LAN) here uses Thin Wire Ethernet and consists of 4 personal computers, and one of them serves as the "server" of the system. For the way of interconnecting all the computers in the network system, the bus topology is employed, in which transmission of data from any station on the network can be received by all other stations at a time. Protocols

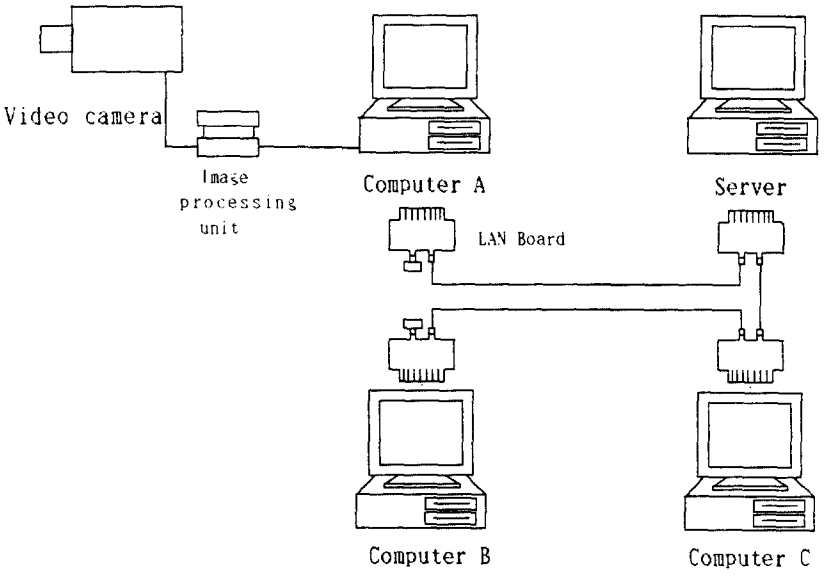


Fig.1 Schematic of system

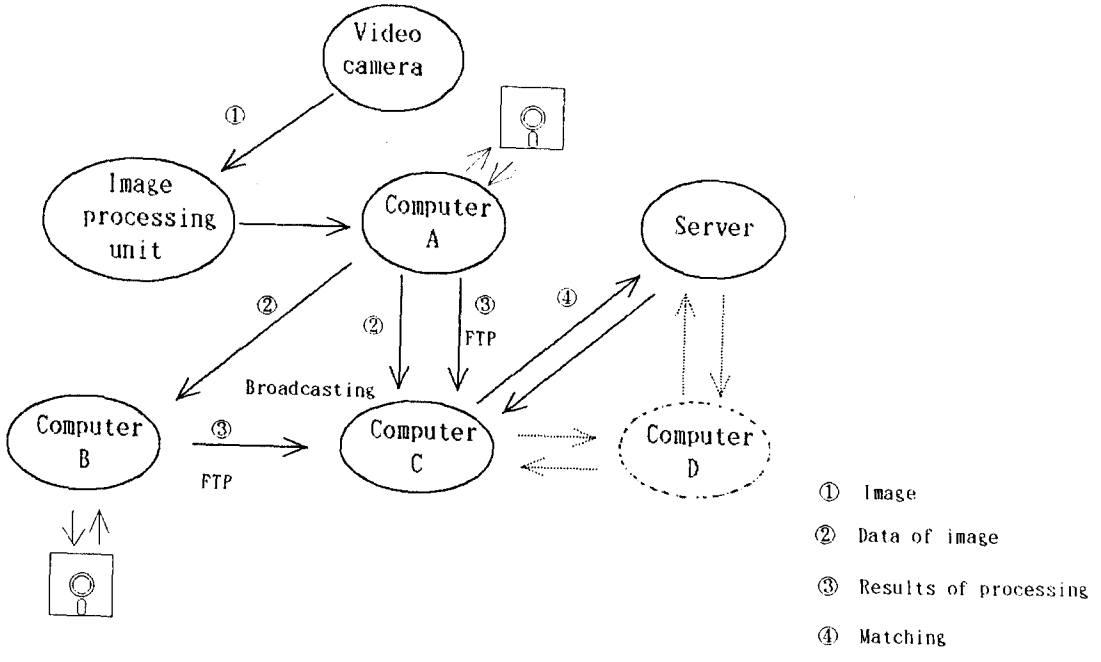


Fig.2 Data stream

implemented are Transmission Control Protocol(TCP) and User Datagram Protocol(UDP) in the middle layer, and Internet protocol(IP) in the lower layer. Hence the present system supports a variety of operation as will be mentioned.

Data as visual information are figures and letters written by hand and are inputted by using a video camera together with image processing devices. Here letters are restricted to a kind of Japanese characters "katakana."

2)Data Flow in the Network

The data flow on the network is shown in Fig.2. A gray scale threshold is set in displaying figures and characters on the CRT. The gray level of visual information inputted through video camera, at each pixel, is compared with a prescribed value of threshold when the image is depicted on the CRT of Station A. If the gray scale level is lower than the prescribed level, then it indicates that the particular pixel is relatively dark. This sort of information is transmitted to Stations B and C together with the information about their coordinates by using the function of UDP.

At Stations A and B, figure processing and letter processing are made in parallel in order to check if the particular pattern inputted includes figures and/or letters. A superordinate protocol, File Transfer Protocol(FTP), of TCP supports file transfer between the clients on the network. The results of the analyses in Stations A and B are transmitted to Station C by using this function of the FTP. Inside the computer C this turn, the contents sent from those stations are checked the possibility of identifying those data as a hollistic whole consisting of more than one object such as a scene (=a set of concrete (sometimes absolute)objects) or a word (=a set of characters). For this purpose, utilized are the pieces of past experience stored in the the auxiliary memory of Server. This way of memorizing experience is effective in applying to the extended system which is a combination of the present system and the system for verbal aspect of cognitive activity [4].

The content of each stage of processing will be described below. Brief description in the parentheses right after each process' name indicates how the data is dealt with in that particular process.

i) Process 1
(Station A to stations B and C)
At this stage of processing, the data stored in the station A at this moment is transmitted to the other computers on the network through the bus ("broadcasting").

ii) Process 2
(Individual procedures inside Stations A and B, and the ones from A to C and from B to C)
Here in this stage, concrete objects (characters and plane figures) are categorized into characters and figures, and in this sense, they are given rough interpretation. Inputted information, as it is, is not checked yet at this moment if it includes characters or figures or both. While in the case of figures, outer configuration or inner configuration as a contour depicted with a thick line identifies what the figure is on the CRT, in the case of characters instead, each of both configurations has its own role and hence both are required in identifying the character. So the parallel execution of both figure and character processings by using two computers makes it possible to identify any visual information (characters and plane figures).

Let us suppose that the inputted information includes n pieces of objects each of which may be a plane figure (a contour) or a character. And let the set of those objects be denoted J:

$$J = \{j_i ; i=1,2,\ldots,n\}.$$

The processes performed in Stations A and B can be expressed by some correspondences from the set J to the set of characters and the set of (contours of) plane figures, respectively. Let the sets of characters and plane figures be the set

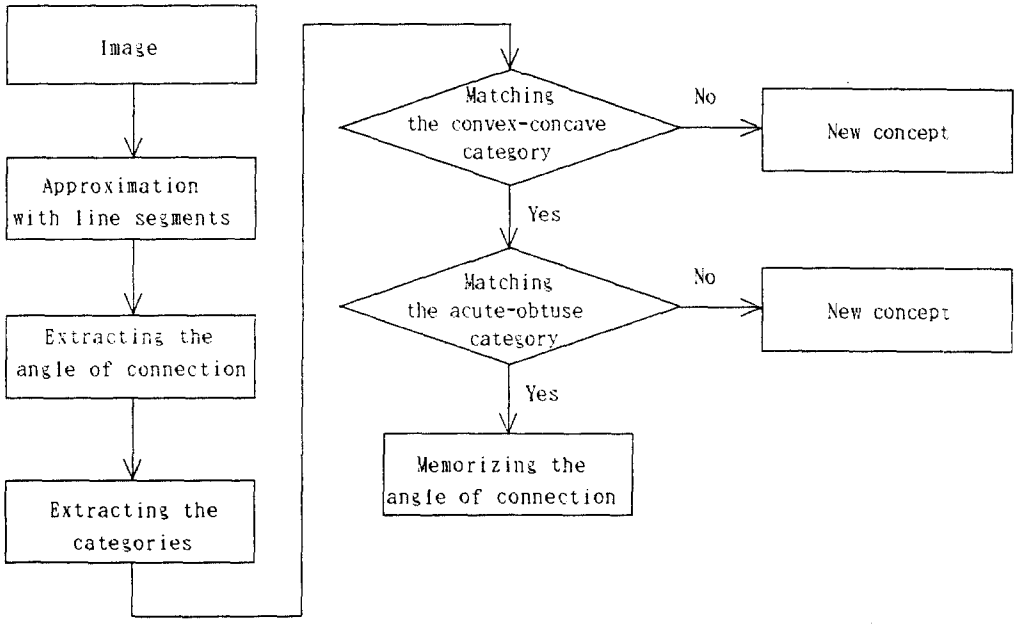


Fig.3 Flowchart of processing in perception level

A and the set B, respectively. And let these correspondeses be denoted as f_a and f_b, respectively. Usually f_a and f_b are considered to be functions. But sometimes there is a possibility of f_a's having more than two objects (characters, in this case) to correspond to a single object in the set J. That is, f_a may be "one to many." This may also be the case in the correspondence f_b. In such cases, f_a and f_b should be called relations, instead of functions. The notations for these correspondences will be the ones for functions, the basic idea of f_a and f_b still including the possibility of relations. Then

$f_a : J \rightarrow A$, (processing in Station A),

$f_b : J \rightarrow B$, (processing in Station B),

where the sets A and B are subsets of characters and plane figures, respectively. The function f_a makes j_i correspond to some element a_i of the set of charactersA if j_i satisfies the necessary characteristics for a character. And if the object j_i does not seem to be any character, then f_a makes it correpond to nothing. And instead the function f_b finds the possibility of a correspondence of j_i to some element of the set B. In the case of the example,

$J=\{j_1, j_2, j_3, j_4, j_5\}$,

$A=\{a_1, a_2, a_3\}$, $B=\{b_3, b_4, b_5\}$

where the sets A and B indicate that only j_3 is recognizable as both a character and a figure, though the other objects j_1, j_2, j_4, and j_5 are recognizable as either a character or a figure and not as both. The details will be described in (iii)Process 3. The algorithm taken in the processings f_a, f_b are like the flowchart in Fig.3.[2]

To understand the inputted information as a shape in the level of perception, the information should be symbolized in some sense. Here, visual information which only includes characters and plane figures (as contours) in it is sectionally linearized and is recognized as sequences of line segments with the knowledge of connection relations of line segments and of the shape's feature.[3] As is already mentioned, the outer configuration and inner configuration

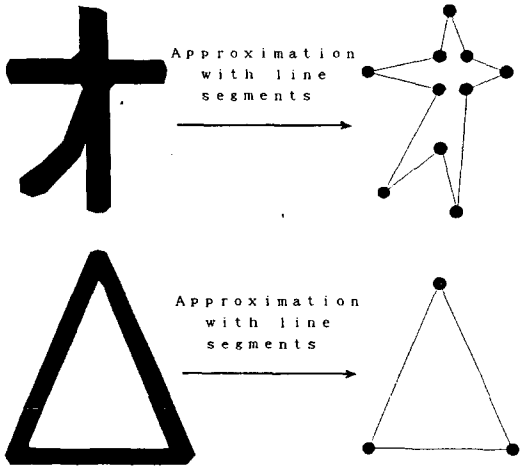


Fig.4 Approximation with line segments for a letter and a figure

have different meanings. So the configuration(s) dealt with in sectional linearization depends on which is the objective of the processing. For figure identification, the inner configuration is used. For character processing, both cofigurations are employed. Examples are given in Fig.4. For feature extraction through sectional linearization, employed are the ideas of convexity-concavity, acuteness-rectangle-optuseness in line segment's connection relation, and more accurate connection relations if necessary. The connection angle itself for every succesive line segments is the last means to judge if an object belongs to a certain category (concept) when other rougher identification does not work well. Fig.5 shows an example of utilizing the convex-concave relation (relation R in the figure;The symbols + and - denote convex and concave relations, respectively.) and the acute-obtuse relation (relation P in the figure; Rectangle is also included, and 1, 0, and -1 are employed to indicate "acute," "rectangle," and "obtuse," respectively.). Objects on the CRT are categorized by these classification and utilized in recognition. And if either one of the

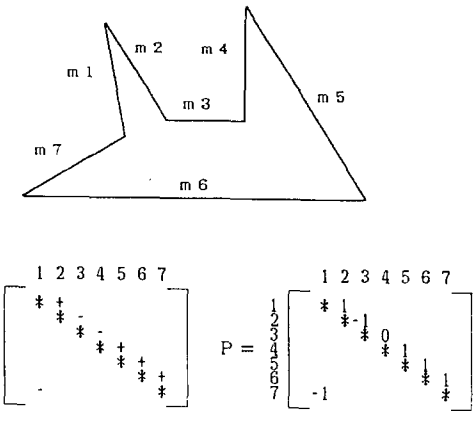


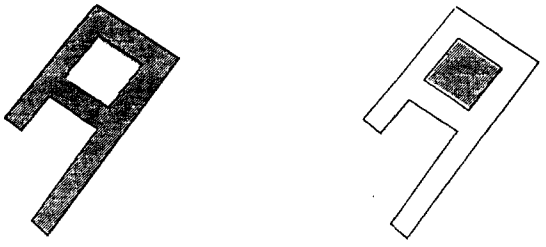
Fig.5 Figure and its relation matrices

categorizations by those relations does not hold for a new object inputted, then that information is used to generate a new categorization, as the maintenance of those relations by an object indicates that the particular object belongs to a category established through the past experiencing.[2] This sort of information is stored in memory and utilized in the future better recognition as part of the past experience. For this purpose, the idea of experience sequence [1] is applied. In this stage, as is mentioned above, objects have possibility of being recognizable as both a character and figure. This possibility will be reduced in the next stage, Process (iii). But this possibility can not be perfectly excluded, since there are such ambiguous objects in the real world.

iii) Process 3

(Procedure executed inside Station C and the one from C to Server)

This part corresponds to the highest level as human cognitive procedure. The outputs from the procedures inside Stations A and B, including the contents of the sets Aand B, are sent to C for this stage of Process (iii). Fig.6 shows one of the typical differences between the plane figure processing and letter processing in the



Letter processing　　Figure processing

Fig.6 Letter and Figure processing of a katakana "タ"

case of a katakana "タ." As is already mentioned, the figure processing employs inner configuration. So, for this case, the figure processing results in misunderstanding of contour, while the letter processing is successful in sensing the shape of the character "タ." As this example shows, the difference can be found in the difference in the occupied gray area between the two processings. The inclusion relation can be available indistinguishing figures from characters. In the example given in Process (ii), j_3 is recognizable as both. But comparing the gray area Oc shared by the outcome from the character processing includes that from the figure processing Of may give the conclusion. That is, if Oc includes Of, then the object can be judged to be a character. In the case of having equal Oc and Of, the object has both possibility. So in that case, the knowledge of the context in which the object should belong to is useful.

Discussion and Conclusion

Application of personal computer network to a cognitive system for visual concept understanding is effective in the sense that parallel processing from a variety of viewpoints is essential in examining feasible solutions for recognition. Although, here, figure and letter processing are dealt with, the present method will be more effective in its application to bilateral translation between image and languge which is under development by the authors. A systematic grasp of a sequence of characters as a "word" or of a set of objects as a "scene" is effective in processing characters and plane figures in parallel. And it is also necessary in the bilateral interpretation of concepts mentioned above. To be developed in order to establish more effective and intelligent systems is the introduction of the idea of context[4], a sophisticated methodology for reasoning and recognizing concepts by using effectively the advantages of image and language, and so on.

Acknowledgment

This work is partly supported by Japanese Government subsidy for aiding scientific researches. The authors greatly appreciate this support.

References

[1]Y.Sakai, "The idea of Experience Sequence for Man-Machine Interfaces," Proc. of Int. Conf. on Manufacturing Systems and Environment, May28-June1, Tokyo, pp.319-323, 1990.

[2]Y.Sakai, and T.Murahashi, "Experience Based Realization of Conceptual Functions for Man-Machine Interfaces," Proc. of IECON'91, Kobe, Oct.28-Nov.1, pp.1605-1610, 1991

[3]M.Nagao, "Gazou-Ninshiki-Ron (Theory of Image Recognition)," Korona, 1983.

[4]Y.Sakai, and K.Nomura, "The Verbal Aspect of Concept Understanding," '92KACC, Seoul, Oct.19-21, 1992.