IMES: Integrated Media Environment System

Toshiyuki SAKAI (Kyoto Univ.)

The IMES project at the department of information science at Kyoto University was started officially in 1985 as a 3 year project under the sponsership of the Ministry of Education and Culture, Japan. A general concepts of the IMES system is shown in the following:

Integration of

- Information Media: Codes, Speech,
  Images(document, line drawing, photographs)
- Transmission Lines: Paired line, Coaxial cables,
  Optical fibers
  LAN : Ethernet, Omninet, Simplenet,
- Processing Machines : Personal computers(terminals),
  Work-stations, Host computers

Actual network are shown in Fig.1.

Front-end network is for users and any terminals may connect to any host computers (in back-bone network) and workstations or super-minicomputers on I/O·WS network at a same time in the form of 1:n connection.

IMES is able to handle coded signal, speech sound and various kinds of static images & dynamic images in real time. Amount of data differes so widely according to the representing media of information as follows:

coded characters on 1 page (1800 characters)	28.8KByte
bi-level image	4064KByte
Modified Huffman code of the above	800KByte
grey-level image	8x4064KByte
color photograph	3x8x4064KByte
synthesized speech for reproducing 1 page sentence	～4200KByte

We have developed various kinds of hardwares, networks and related softwares in order to realize research targets related to the most fundamental and practical application fields of information science.

The following examples are executed within a period of several seconds or several minutes.

[S/S' system]

The important key point to ascertain in terms of determining what kind of information processing to undertake is that of knowing whether the object pattern is generated in an S/S' system or not. In the case of a multiple of images on one sheet, if there exist a restriction or partition rule, then it may be possible for the machine to process it as an S/S' system.

As can be often seen in some types of magazines, photographs of natural scenes on which character strings are superimposed without regular borderlines or, illustrations in which non-character figures and characters are mixed in, are two complicated examples of S/S' systems. When partitioning restrictions are maintained, it does not matter for machine processing whether the constituent elements are either characters, text lines, graphs, tables, illustrations or photographs and so forth.

[Hierarchical Structure]

Hierarchical structure of real world patterns provides for the systematic and compact descriptition of patterns, so that we can indicate specific layer (level) of description most appropriate to the processing purpose at communication stage.

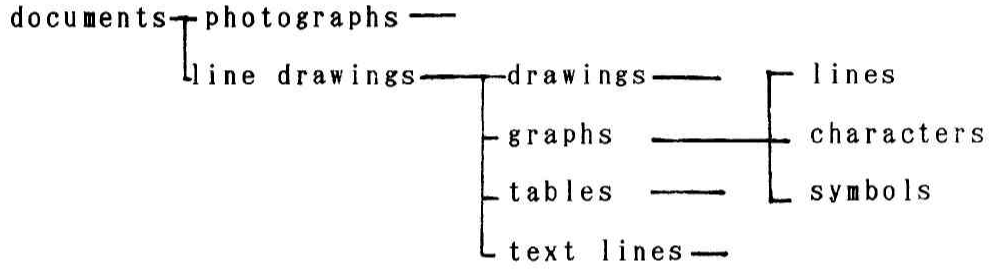
Hierarchical structure is analogous to natural language grammar in which there is a hierarchy of symbols, words, phrases, partial sentences, sentences and then finally the meanings itself. Depending on

the generation rules (for example, transformational grammar) or parsing rules, grammatical tree structures are expandable or reduceable to the leaf side or the root side.

Real world pattern have a tree structure in which the nodes correspond to the constituent elements of that layer of the real world pattern and/or the model world.

If the hierarchical structure nodes are in good correspondence to the user's(human) concepts or entities, then man machine interaction can occur smoothly and conveniently to both sides of the interaction.

A simple examples of hierarchical structures of documents is shown in the followings.



[Adaptive Control]

Adaptive control is that type of control which is adaptable to the signal or data as it is derived from the input pattern under some definite rules.

Data-driven adaptive control is easily found by utilizing binarization threshold value determination in which signal and background discrimination are made by using the statistics of the grey-level of pixels(in the case of line drawings, small occurrences correspond to the signal while majority occurrences are background). A somewhat different example is the quality descrimination of documents by measuring the occurrence of certain legal patterns.

Other examples of adaptive control are the descrimination of photograph areas or line drawing areas as decided by either the line

model or the photograph model (non-line model) in the bottom-up stage. Region or text line descrimination by the mesh method is also a type of adaptive control based on horizontal (and/or vertical) projection profiles (in other words, histograms of black pixels in a cretain mesh area).

[Intelligent Processing]

Intelligent processing may be defined to be rules or models in the machine, in which: 1.) procedures of selection are executed among various possible candidates, 2.) understanding process continues by consulting model or rules using a method of matching or evaluation in multi-dimensional space.

Intelligent systems inevitably perform misoperations or misjudgements making correction operations or backtracking mechanisms necessary. If at least one part of the processing satsifies the conditions of the above mentioned definition, then it is intelligent processing. The followings can be considered to be examples.

(i) In line drawings, a line segment has the possibility of being a stroke of one character or a symbol, a connecting line of a drawings or a line making up a table or graph. According to the understanding process it is decided as being one or the other by applying various kinds of rules depending on the environment.

(ii) In maps, a character is decided by the rules of the exclusive characteristic properties. Determing to what kind of character string it may belong, however, is not so simple, making this processing also, an example of intelligent processing.

[EXAMPLE 1] Intelligent Processing, Storage and Adaptive Editing of Text Documents

On one and the same document sheet, there may be text lines,

photographs, illustrations, graphs, tables and drawings, etc.. Text documents may be divided into head line (title) characters, abstract, text lines and other sections including author name, affiliations and so on. As a man-made document, it can be said to have a definite kind of format. As a basic conceptual tool, rectangular layout regions are considered to be fundamental. For a text-document, the basic unit is the text line which in various arrangements makes up the text document.

It can be easily seen that character strings are made up of combinations of isolated (alphabetic) characters each in an appropriately sized rectangle. Text lines are specified by their length, width, and line spacing no matter whether we considering a vertical or horizontal line arrangement.

For document understanding, no matter what the input method, and no matter how different the input pattern quality may be, the text line should be recognized automaticaly, adaptively , with high accuracy.

The large size block units which are formed by the accumulation of text lines in the same column, are labeled with a sequence number which is in accordance with the human reading sequence (from top to down, and from right to left for Japanese newspapers).

Text document understanding has proved to be practical not only for Japanese newspapers but also for English newspapers without the machine needing to be changed. The machine scale, processing speed and cost can also be said to be practical from various points of view including that of storage, re-editing, coding and distribution to distant terminals through networks.

An adaptive, re-edited example is made using the text line as a unit for re-arrangement. If the reader wants to see an enlarged character of a specific head line, abstract or text line, it is possible to output this in the specified form. The size of the head lines in newspapers is large enough so that it is possible to use

OCR(optical character reader) to convert the character image into a corresponding code. This makes it easy for humans to retrieve articles without additional human pre-processing. Because a machine knows every data item of articles, a database formed after this pattern understanding processing, may deliver keywords from the head lines and keywords from sentences in the abstract.

[EXAMPLE 2] The Transformation and Translation of Maps and Drawings by the Extraction of Characters and Symbols

(1) White map generation by the extraction of characters and symbols

The most typical examples of S/S' systems are maps which consist of land shape figures which are superimposed upon by characters under some restrictions in terms of position and size. Depending upon the purpose of the map, priority is given to either the figures or to the characters resulting in the lower priority type of information at times being hidden. Some times, the figures and the characters are superimposed independently without need of considering the other information's situation.

Intrinsically, after extracting the characters and symbols, it is necessary to restore the background previously hidden (figure) parts. Also, on the other hand, we want to be able to add new character information and symbols, illustrations, enlargements, etc., to the white map as per human instruction.

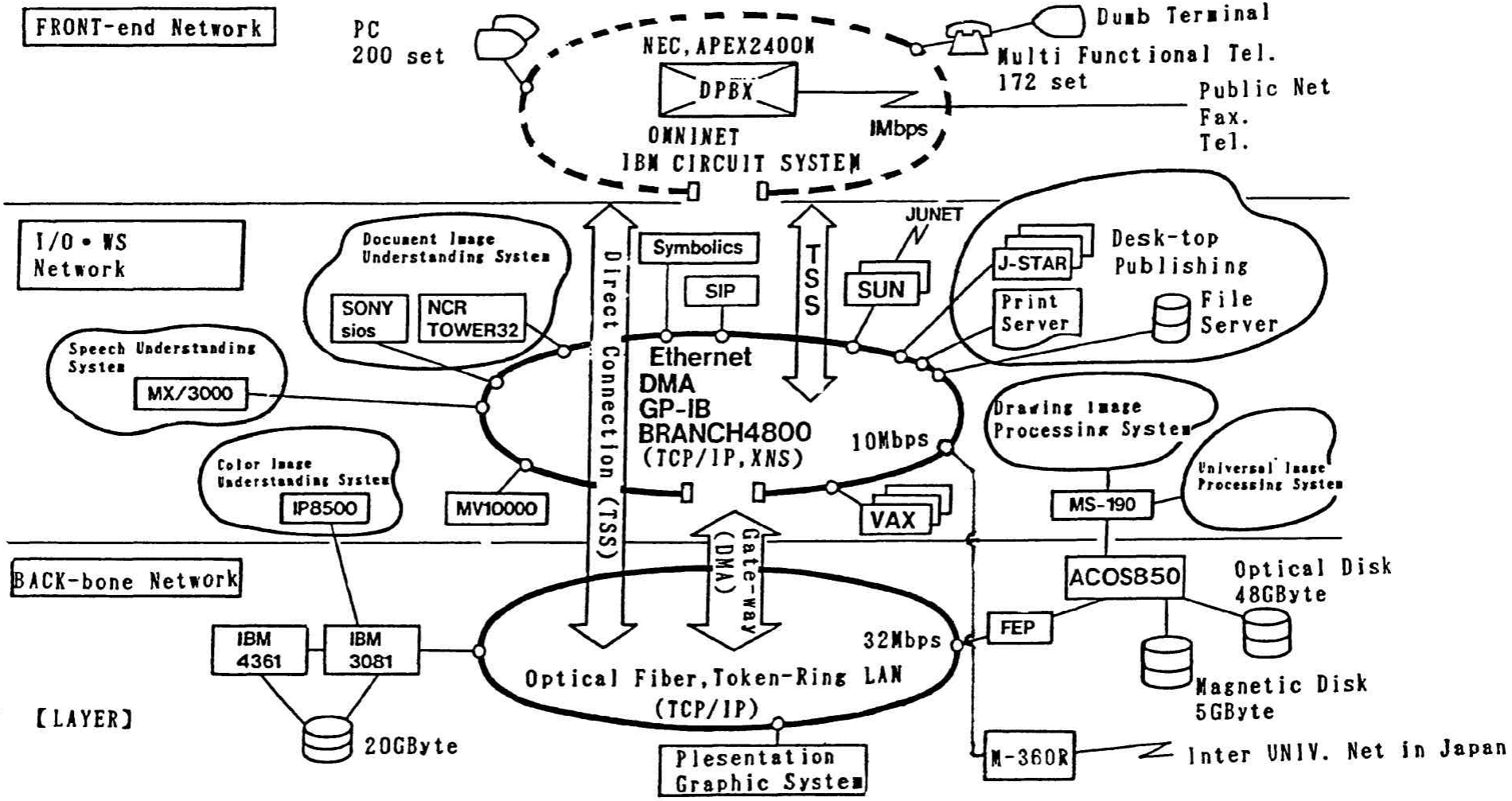
In general, character elimination in white map processing is not sufficiently good, so post-editing by the human is necessary for achieving satisfactory results. In terms of the machine processing itself, the lines of the land shape figures are long in length as compared to their width, while characters have complex black pixel structures within isolated circumscribing rectangles.

(2) Intelligent adaptive transformation and translation of line drawings which maintain character quality and line drawing line widths.

The extraction of characters and symbols, and line descrimination is the same as in (1), but in this case we also consider characteristic points. Characteristic points make up the constituent elements of line drawings. One example of the high level (syntactic hierarchy) description of line drawings employs as its constituents, blocks containing character strings, connecting lines and characteristic points.

The original drawing is analyzed into three constituent elements at the syntax level and we attempt to understood these relations from a topological point of view. Then the machine transforms the image into the specified size (the horizontal and vertical scales need not be the same) or translates the original character strings into other types of character strings. This translation may be natural language translation or the translation of abbreviations into normal words or vice versa. In these cases, the character string no longer has the same length so the surrounding block must be adaptively re-determined.

Intelligent processing is also necessary to insure the beauty of the drawing by maintaining the regularity of the block size, an adequate relative size of the characters and so on.


FRONT-end Network
PC
200 set
NEC, APEX2400M
DPBX
OMNINET
IBM CIRCUIT SYSTEM
1Mbps
Dumb Terminal
Multi Functional Tel.
172 set
Public Net
Fax.
Tel.
I/O • WS
Network
Document Image
Understanding System
SONY
sios
NCR
TOWER32
Symbolics
SIP
Direct Connection (TSS)
TSS
JUNET
SUN
J-STAR
Desk-top
Publishing
Print
Server
File
Server
Speech Understanding
System
MX/3000
Ethernet
DMA
GP-IB
BRANCH4800
(TCP/IP, XNS)
10Mbps
Drawing Image
Processing System
Universal Image
Processing System
Color Image
Understanding System
IP8500
MV10000
VAX
MS-190
BACK-bone Network
Gate-way
(DMA)
ACOS850
Optical Disk
48GByte
IBM
4361
IBM
3081
32Mbps
FEP
Optical Fiber, Token-Ring LAN
(TCP/IP)
Magnetic Disk
5GByte
【LAYER】
20GByte
Plesentation
Graphic System
M-360R
Inter UNIV. Net in Japan