A STEPWISE RECOGNITION METHOD OF LIBRARY CATALOGING CARDS ON THE BASIS OF VARIOUS KINDS OF KNOWLEDGE

Toyohide Watanabe
Qin Luo
Yuuji Yoshida
Yasuyoshi Inagaki

Reprinted from PROCEEDINGS OF THE TENTH ANNUAL INTERNATIONAL PHOENIX CONFERENCE ON COMPUTERS AND COMMUNICATIONS, Scottsdale, Arizona, March 27-30, 1991

A Stepwise Recognition Method of Library Cataloging Cards on the Basis of Various Kinds of Knowledge

Toyohide WATANABE, Qin LUO, Yuuji YOSHIDA and Yasuyoshi INAGAKI

Department of Information Engineering,
Faculty of Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-01, JAPAN

***** Abstract *****

It is one of important issues to extract the meaningful data from library cataloging cards automatically and then construct them as a library database, because each library has accommodated a great deal of cataloging cards. The traditional methods based on the character recognition technique can be partly applied to this issue. However, these methods are not always successful for the cataloging cards, specified by different layout structures and various description rules.

This paper proposes an experimental method to extract and classify the data items automatically from cataloging cards. Our basic idea for the recognition strategy is to utilize various kinds of knowledge stepwisely on the basis of the generation/verification process of object hypotheses: in our approach, the knowledge such as layout information of card structures, relationship information among cataloging items, and property information of data items is effectual.

1. INTRODUCTION

The issue about the document understanding is an important research subject, today[1]. The main research objective is to extract and classify the meaningful information automatically. Traditionally, the methods to extract characters from some formatted texts have been investigated as an application of the OCR or character recognition techniques[2]. These methods are not successful to classify distinct data items even though they are adaptable to the character recognition. Additionally, these methods can not always apply to documents with various types of layout structures and/or various kinds of description rules because the recognition process is controlled only on the basis of characteristic values analyzed statistically from the sample data. More excellent methods are too strongly required[3].

Such a subject is never excluded even in the library managements[4]. In the libraries, a great deal of library cataloging cards have been composed to index individual books easily and manage them effectually. However, the composition of machine-readable cataloging data is one of difficult problems in many libraries. This is because the cataloging cards are not designed so that computers can recognize the card forms directly. Thus, the traditional approaches are not always applicable to the recognition of cataloging cards. Namely, the traditional approaches are not excellent solutions for the next problems:

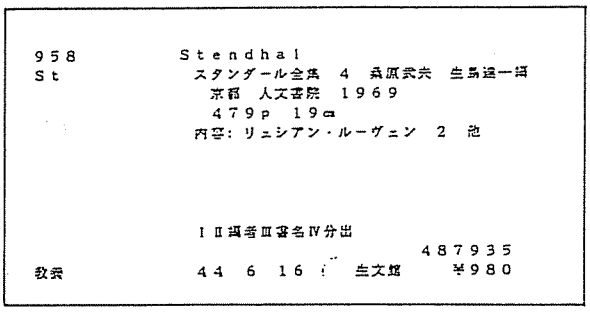
1) The data items in individual cataloging cards are not only different in accordance with each book, but also some of them may be omitted;
2) The data items in individual cataloging cards are, in many cases, not always assigned to the same positions;
3) The layout structures or description rules have been more or less changed in each library until today;
4) The characters typed in cataloging cards are blurred and indistinct because users have manipulated them directly.

In this paper, we propose an experimental method to extract individual data items from the cataloging cards and classify them into the cataloging item classes automatically. Our approach, characterized as the top-down method, is to recognize the data items on the basis of various kinds of knowledge, concerning the cataloging cards[5]. The knowledge is not represented as the same-level information, but is mutually specified with the hierarchical relationship. Namely, lower-level knowledge is successful to recognize the processing objects interpreted by higher-level one.

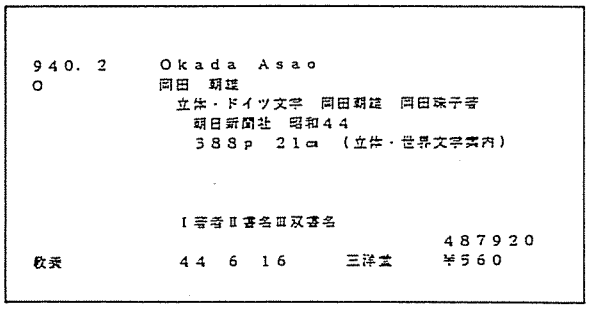
2. CHARACTERISTICS OF LIBRARY CATALOGING CARDS

In library cataloging cards, various cataloging data are composed under the layout structures of cataloging cards, the mutual relationships among data items, and the particular properties of individual data items. Such description methods are commonly based on the generally authorized cataloging rules[6]: concerning the catalogs of Japanese books, NCR(Nippon Cataloging Rule)[7] is used as the standard one. However, the standard rule is rarely applicable to each library, as it is. For example, consider the cataloging cards of Japanese books in our University Library. 3 types of cataloging cards are available as shown in Fig.1. These cataloging cards are possibly distinguished owing to the geometric and syntactic properties. Of course, many varied cataloging cards are in existence, corresponding to the length of data or the number of multiple data as well as the abbreviation of data items.

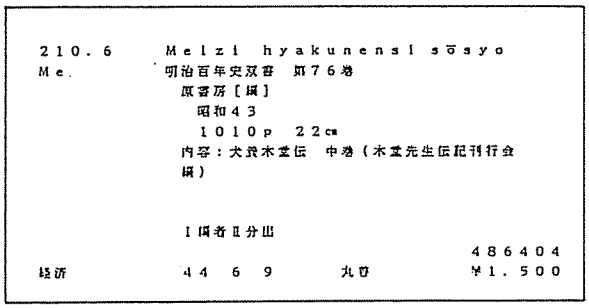
Generally, we can classify the compositive data items with respect to their mutual relationships and properties. Fig.2 shows the logical structure. From a spatial point of view, we can distinguish 4 information sections: the call-number section in the left-upper area; the location section in the left-lower area; the bibliography section in the right-upper area; and the management section in the right-lower area. The distinction of 3 types of our cataloging cards depends mainly on the descriptive relationships among



(a) type A: translated book



(b) type B: original-writer book



(c) type C: series book

Fig.1 Types of Japanese cataloging cards

data items to be located in the bibliography section. Additionally, individual character lines in the bibliography section are specified under the geometric relationships such as one-character left(or right)-shifted forms between the upper and lower lines. Some data items accompany with the special keywords such as "著" for the author item, "訳" for the translator item, "版" for the edition item and so on.

It is effective to make use of such descriptive knowledge in case that we attach to the subject for the automatic identification of cataloging items. If we can conceptualize various kinds of knowledge, it is not so difficult to interpret various cataloging cards. The approaches based on the character recognition techniques are not applicable to various types of cataloging cards besides a similar card class. Moreover, these approaches can not classify individual character strings into the meaningful cataloging items even though they could extract all characters successfully from the cataloging cards.

3. FRAMEWORK OF KNOWLEDGE-BASED APPROACH

Our knowledge-based approach is a top-down method[8].

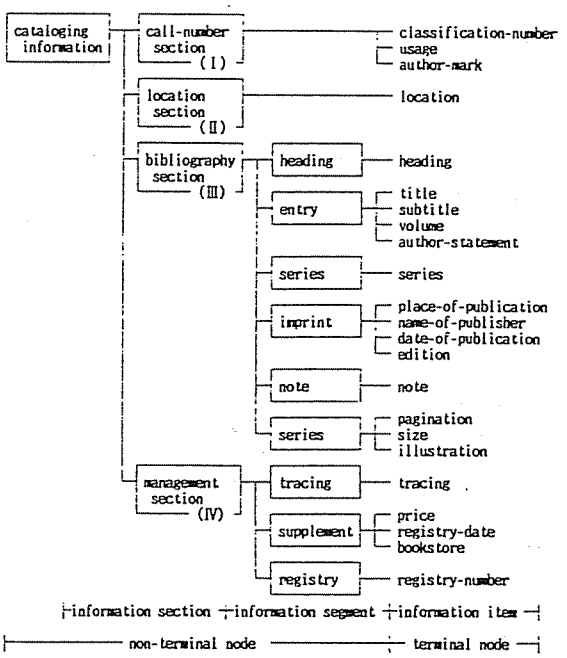


Fig.2 Logical structure of Japanese cataloging cards

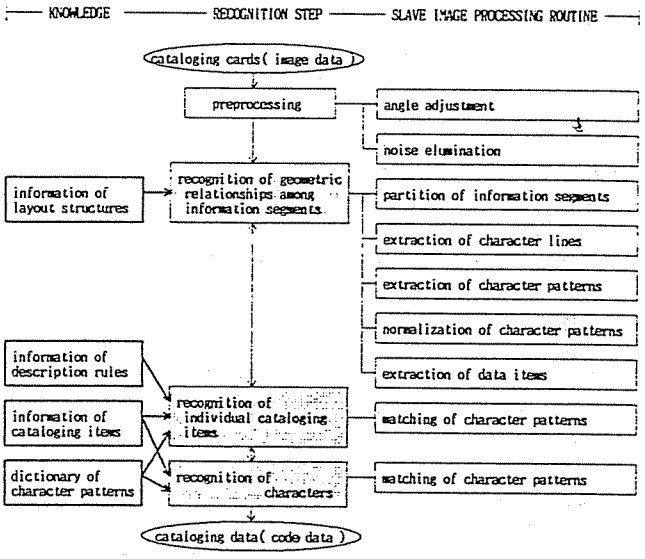


Fig.3 Processing flow

Image processing routines are invoked as slave routines under the recognition procedures. The previously recognized image objects become source data for the following recognition procedure. Such a processing flow is shown in Fig.3. The identification mechanism is divided mainly into 3 recognition procedures: recognition of geometric relationships among information segments on the basis of the layout structures; recognition of individual cataloging items with respect to the description rules, item properties and character pattern dictionary; and recognition of characters with the information of item properties in addition to the traditional character pattern dictionary. These 3 types of

recognition procedures interact mutually: the forward link transfers the cataloging image objects, which interpreted by the higher-level recognition procedure, into the lower-level recognition procedure in Fig.3; and the backward link indicates to be retried again from the higher-level recognition procedure in case that the lower-level recognition procedure found the inconsistency among descriminated objects or could not interpret the characterized objects in detail.

Fig.4 represents the hierarchical relationship among various kinds of knowledge, concerning the cataloging cards. 3 kinds of knowledge is applicable to 3 types of recognition procedures. The information of layout structures deals with the spatial/geometric relationship among information sections/segments from 2-dimensional point of view. The information of description rules indicates the constructive/combinative relationship among information items, which is adaptable to individual information segments/sections, from 1-dimensional point of view. Finally, the information of cataloging items points out the characteristics of information items with respect to their own discrimitive properties.

Our approach can provide excellent solutions for the problems 1)-4). As for the problem 1), our approach is successful because our description rules can specify the logical relationships among data items without depending on individual cataloging cards. For example, the recognition procedure can identify the particular cataloging items by comparing the description rules with the information about neighboring data items even if any cataloging items were omitted. As for 2), it is possible to pick up the same class of cataloging items on the basis of the knowledge about both the spatial relationships among information sections/segments and the neighboring relationships among information items. In our approach, the positional information is not specified in advance, but is determined stepwisely by the recognition procedures: the recognition procedure interprets image objects by using only the logical information. As for 3), our approach can interpret varied cataloging forms/formats. Namely, it is easy to add(or replace by) new knowledge even if the layout structures or description rules were altered, because the knowledge is managed declaratively. Finally, concerning the problem 4), our approach is powerful in comparison with the traditional character recognition technique. This is partly because our approach can distinguish the noises from characters by means of compensatory knowledge about spatial/geometric and constructive/combinative relationships among individual cataloging items, and partly because the knowledge about cataloging items can help selecting exactly the particular characters from the candidates.

4. KNOWLEDGE ABOUT LAYOUT STRUCTURE

Our knowledge about the layout structures is conceptual information derived from spatial, topological and geometric relationships among information segments/sections. The conceptual information to be independent of the particular instances of cataloging cards is successfully applied through the interpretative recognition strategy.

The basic strategy in our approach is to interpret the

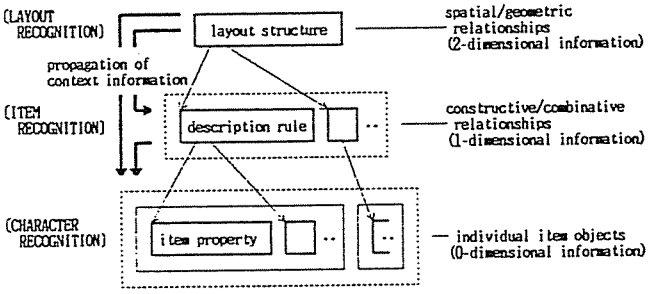


Fig.4 Hierarchical structure of knowledge concerning library cataloging cards

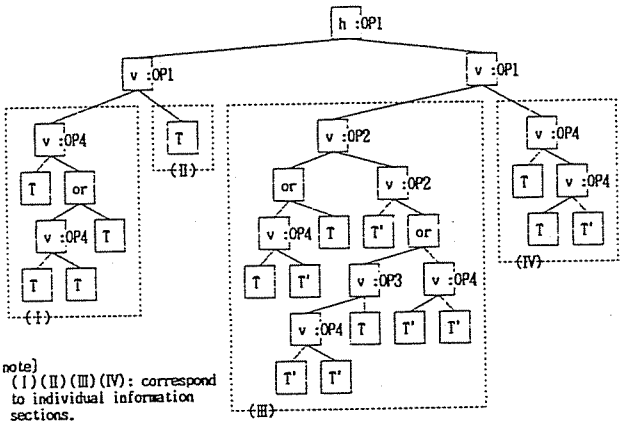


Fig.5 Tree representation of layout structure

layout structure attended logically to individual instances on the basis of the fact that the information sections/segments construct neighboring relationships in the vertical and horizontal directions spatially. We adopt a hierarchical representation structure(binary tree structure) to specify these neighboring relationships. In the tree structure, the nodes are the hierarchically distinguished blocks, and the branches represent the neighboring relationships among these blocks. The terminal nodes indicate individual blocks, corresponding to information segments/items. For example, the Japanese cataloging cards in Fig.1 can be specified completely as shown in Fig.5.

In this representation, 3 types of non-terminal nodes such as "h", "v" and "or", and 2 types of terminal nodes such as "T" and "T'" are used to control the segmentation process. The node "h" indicates that the next segmentation operator should divide the block in the vertical direction, and the node "v" does to be effective in the horizontal direction. Fig.6 illustrates these effects. In this case, the left and right segments separated by the horizontal cutter, or the upper and lower segments separated by the vertical cutter are uniquely ordered: the order is left-to-right for the horizontal and vertical cutters. The node "or" points out that the child nodes in the tree structure are selective. Namely, this node can deal with the variation of layout structures. While, the terminal node "T" indicates that this block can include only one data item,

and the node "T'" does that many data items may be included in this block yet.

In addition, the practical segmentation method is defined by the cutting operator, attended to each non-terminal node, as shown in Fig.5. Namely, "OP1", "OP2", "OP3"... are the cutting operators. Individual cutting operators cut off several meaningful character areas in the block so as to be consistent to mutual relationships among neighboring blocks, as illustrated in Fig.7. 4 kinds of cutting operators work effectively without using the absolute/relative positional values such as the coordinate values, sizes, lengthes and so on. The cutting operator estimates the effectual area sizes of the following blocks by itself and informs the nextcutting operators of them.

The non-terminal nodes are composed of 3-ary cells: (MOD, OP, CO). The field MOD distinguishes the node type such as "h", "v" or "or". Also, OP indicates the cutting operator. CO keeps the size of the block, set by the cutting operator. Therefore, the segmentation process is controlled autonomously from upper nodes to lower ones along the tree representation. While, the terminal nodes are composed of 2-ary cells: (TYPE, CO). The field TYPE represents the terminal node such as "T" or "T'". CO is the same as that in the non-terminal node.

5. KNOWLEDGE ABOUT DESCRIPTION RULES

The knowledge about the mutual relationships among cataloging items is the logical information in accordance with description rules. This knowledge is the complementary information for the knowledge about layout structures. Namely, the knowledge about layout structures can not distinguish individual data items completely: the terminal node "T'" can not be divided into more primitive cataloging items as illustrated in Fig.5. This knowledge specifies descriptively the constructive relationships among data items in the blocks. For example, the description rule for the entry information segment(e.g."スタンダール全集 4 桑原武夫 生島遼一編" in Fig.1(a)) is specified as follows:

(ex.) title(M), subtitle(O), volume(O), author(O), editor(O), translator(O)

Here, "M" or "O", attended to each data item, points out whether the item is mandatory or optional. "M" means that the data item must not be omitted, and "O" does that it is selective. If this knowledge were applied to processing objects, interpreted by the knowledge about layout structures, together with the knowledge about item properties, individual processing objects can be recognized exactly.

6. KNOWLEDGE ABOUT ITEM PROPERTIES

We can not distinguish individual processing objects by only the knowledge about the mutual relationships among cataloging items, because each data item derived from the description rules is not easily decided without the characteristic properties concerning individual cataloging items. Therefore, the knowledge about item properties is required as the supplementary information for description rules. Additionally, this knowledge is also useful in the character recognition process. We represent this knowledge by using

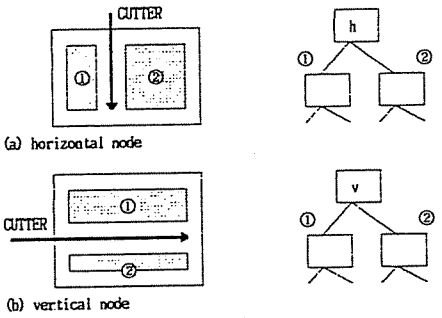


Fig.6 Horizontal segmentation and vertical segmentation

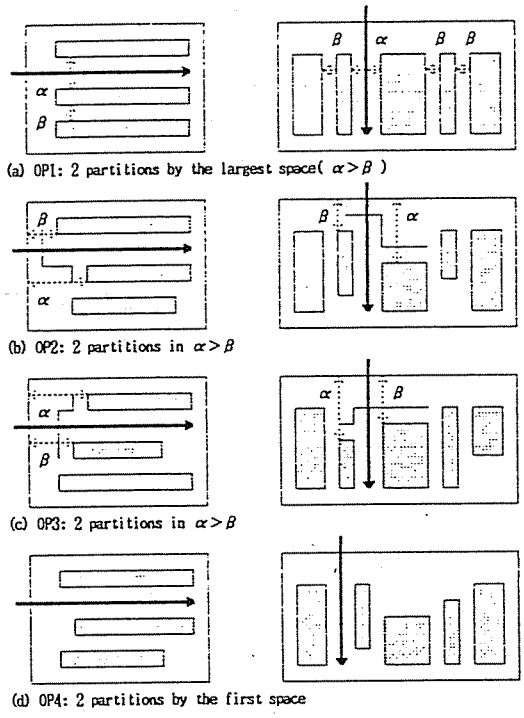


Fig.7 Cutting operators

slot no.(slot attribute)	slot values	decision value
1 (keywords)		
2 (character type)		
3 (number of characters)		
4 (occurrence)		
5 (candidate data values)		

note)
・0 ≤decision value≤3
・character type:
C:kanji, D:number,
E:alphabet, J:kana,
O:others

Fig.8 Slot attributes of item frame

the item frame as shown in Fig.8.

The item frames hold the characteristic properties of individual cataloging items as the slot attributes/values. The keywords, the character set, the number of characters, the occurrence and the candidate data values are preset as the slot values, corresponding to each slot attribute. The decision values associated with every slot indicate the possibility about whether the slot values can decide the processing object as the particular cataloging item. For

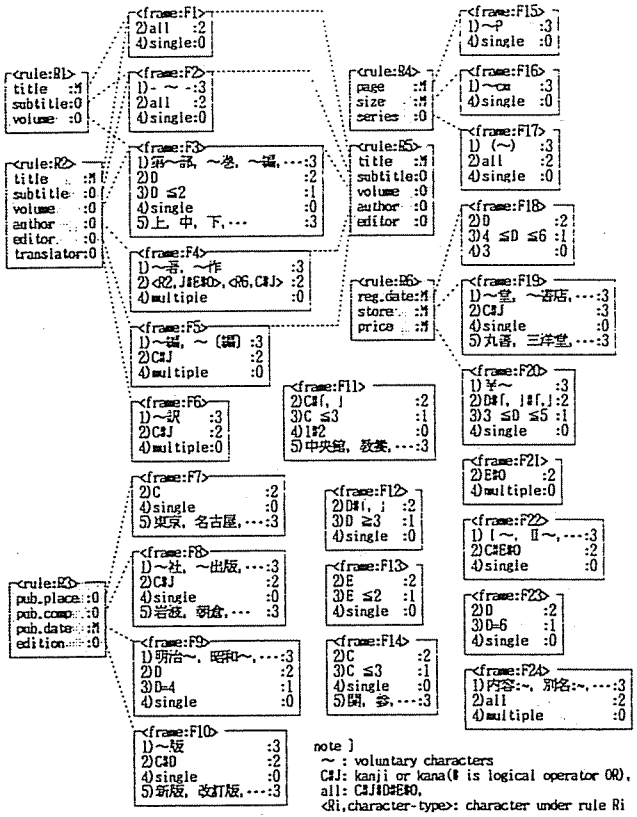


Fig.9 Description rukes and item frames

example, the decision value "3" represents that a processing object can be completely decided as the cataloging item, while "2 or less" points out to be uncertain for the complete decision. This decision process judges by the sum of individual decision values attended to the matched slots as whole: if the sum is 3 or more, then the interpreted processing object is decided as the cataloging item.

Practically, the recognition procedure of data items to be performed under the interpretation of description rules works powerfully with the assistance of the item frames. The relationship between description rules and item frames is arranged in Fig.9. In our experiments, 6 description rules and 24 item frames are effective.

7. EXPERIMENTS AND CONSIDERATIONS

Here, we evaluate our approach through several experiments. The recognition procedure based on spatial/geometric relationships among information segments separates the block in the vertical or horizontal direction. The segmentation result shown in Fig.10 is produced from the original digital image data in Fig.1(a). This result shows not only successful blocks, but also distinguishes candidate data items after extracting each character/symbol and normalizing character sizes. Moreover, we must pay attention that the noises are extracted under the same strategy as the effective extraction of characters.

Next, we show the identified data items under the knowledge about mutual relationships among cataloging items, and

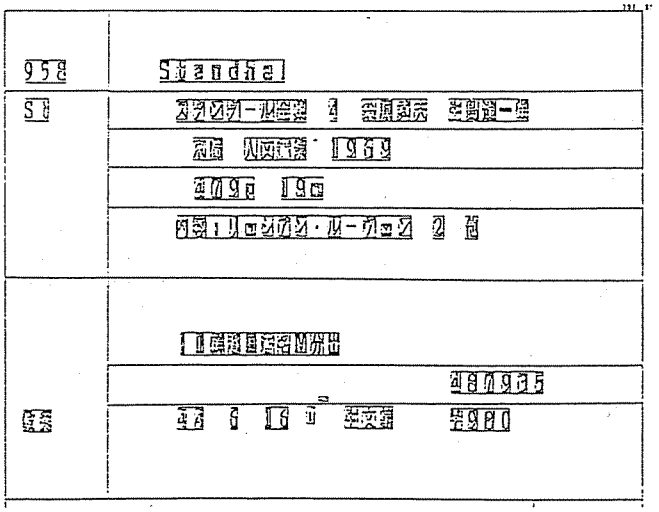


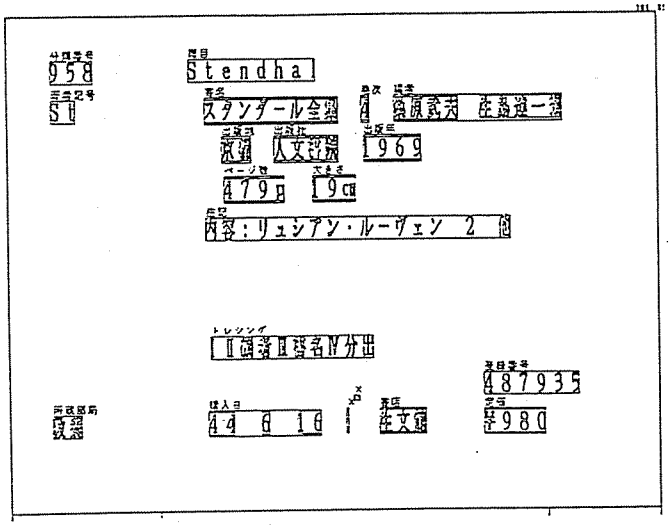
Fig.10 Segmented blocks

the knowledge about item properties. The input data for the recognition procedure of individual cataloging items are the partitioned image data as shown in Fig.10. Some experimental results are arranged in Fig.11. Individual results are shown, corresponding to 3 types of cataloging cards.

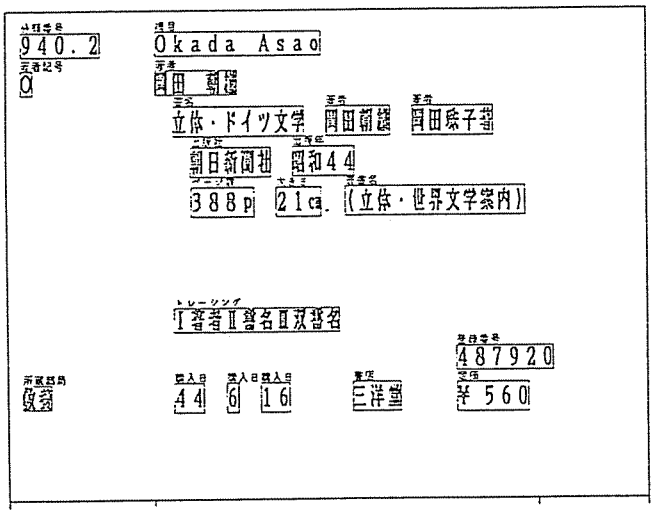
With a view to recognizing documents, the advanced researches on the basis of the knowledge-based approach have been already reported[9-12]. However, these methods are not always successful in comparison with our method. This is because these methods utilize the physical information about layout structures directly. For example, the method in the references 9) and 10) works by the specification of relative positions for each descriptive item, and the method in 11) and 12) works on the basis of the positional information defined approximately about the geometric relationships among individual descriptive items. Thus, these methods are not necessarily adaptable to the transformed documents, whose fundamental structures are irregularly reformed. However, our method is applicable to many variations as shown in Fig.12. The adaptability is mainly derived from our idea based on the spatial and geometric properties among descriptive items. Our principal recognition mechanism is the understanding of spatial relationships among neighboring segments in which individual data items are included. While, the methods in 9)-12) focus on the interpretation of locational relationships among data items directly.

Our approach is adaptable to the cataloging cards with noises. The noises are manipulated as a kind of data items in the segmentation procedure, as shown in Fig.10, but the identification procedure can exclude noises since they are not matchable to the property information of the description rules and item frames, as shown in Fig.11(a). In the traditional approaches based on the character recognition technique the noise elumination procedure is one of the most basic processing modules with a view to making the recognition ratio high.

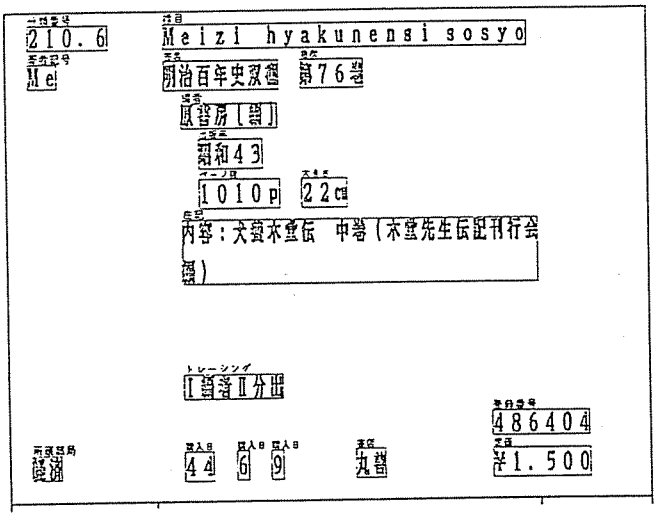
Additionally, our approach is powerful to recognize the blurred and indistinct characters. The cataloging items



(a) type A



(b) type B



(c) type C

Fig.11 Identified cataloging items

and characters are recognized easily in case of being matched to one of candidate data values on the basis of slot attributes of the item frames. For example, the character "編" in the cataloging item トレーシング(tracing) in Fig. 11(b) and (c) will be extracted correctly with the knowledge that the cataloging item is the tracing and the character is one of several restricted characters such as "編", "者","分", "出", "著", "書", "名" and so on. This is because the meaning of each data item can be heuristically determined by the segmentation procedure and identification procedure. Thus, the characters can be effectively recognized on the basis of the corresponding item frame and context information even if the characters are blurred and indistinct.

8. CONCLUSION

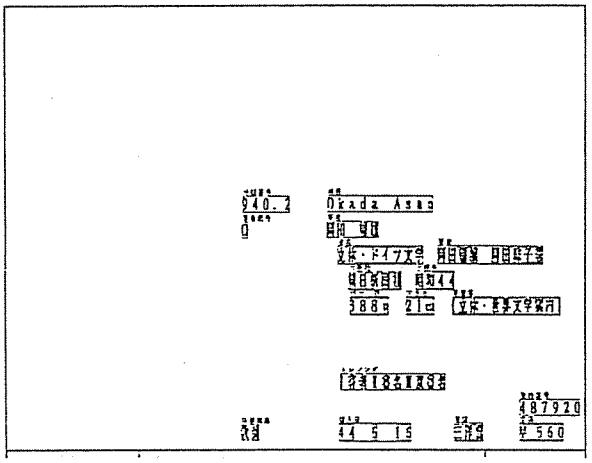
Our approach is not only very powerful in comparison with the traditional approaches, but also successful by comparing with the other similar approaches. In this paper, we investigated the subject about the understanding of library cataloging cards, but its framework is applicable to the other documents with particular layout structures such as pamphlets, letters, office mails, office documents, articles, papers and so on. For example, we can show segmentation results of the other documents in Fig.13.

At least, our approach is more effective with respect to the adaptability, flexibility and applicability. Of course, we must refine our method to be applicable to documents with more complex structures, whose layouts can not be separated simply in the vertical and horizontal lines. Additionally, it is necessary to improve the control mechanism for knowledge in order to make our framework effectual.

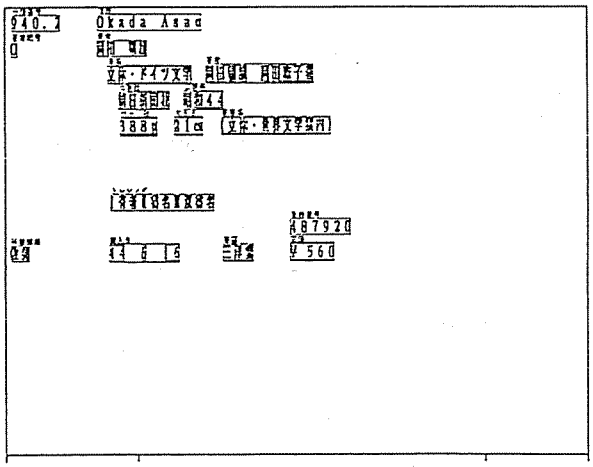
Acknowledgements -- We are grateful to Prof. T.FUKUMURA of Chukyo University and Prof.J.TORIWAKI of Nagoya University for their perspicacious remarks, and also wish to thank the staffs of the University Library in Nagoya University for their many assistances and cooperations. This research was partly supported by the OKAWA Institute of Information and Telecommunication.

References

1) W.DOSTER: "Different Status of a Document's Content on its Way from the Gutenbergian World to the Electronic World", Proc.of 7th ICPR, pp.872-874(1984).
2) C.Y.SUEU:"Character Recognition by Computer and Applications", Handbook of Pattern Recognition and Image Processing(ed.by T.Y.YONG & K.S.FU), P.705, Academic Press Inc. (1986).
3) G.NAGY: "Hierarchical Representation of Optical Scanned Documents", Proc.of 7th ICPR, pp.347-349(1984).
4) H.HASE, M.YONEDA, M.SAKAI & J.YOSHIDA: "On an Automatic Item Classification for Book Cards", EIC trans., Vol.J70-D, No.8, pp.1579-1588(1987) (in Japanese).
5) T.WATANABE, Q.LUO, M.MIZOGAMI, Y.YOSHIDA & Y.INAGAKI: "Automatic Extraction and Classification of Data Items from Library Cataloging Cards by a Knowledge-based Approach", Proc.of MIV-89, pp.67-71(1989).
6) B.S.WYNAR:"Introduction to Cataloging and Classification, 5th edition", Library Science Text Series, P.426, Libraries and Limited Inc.(1976).
7) Japan Library Institute:"Nippon Cataloging Rule", (1977).
8) M.EJIRI: "Knowledge-based Approaches to Practical Image Processing", Proc.of MIV-89, pp.1-8(1989).
9) A.DENGEL & G.BARTH: "High Level Document Analysis Guided

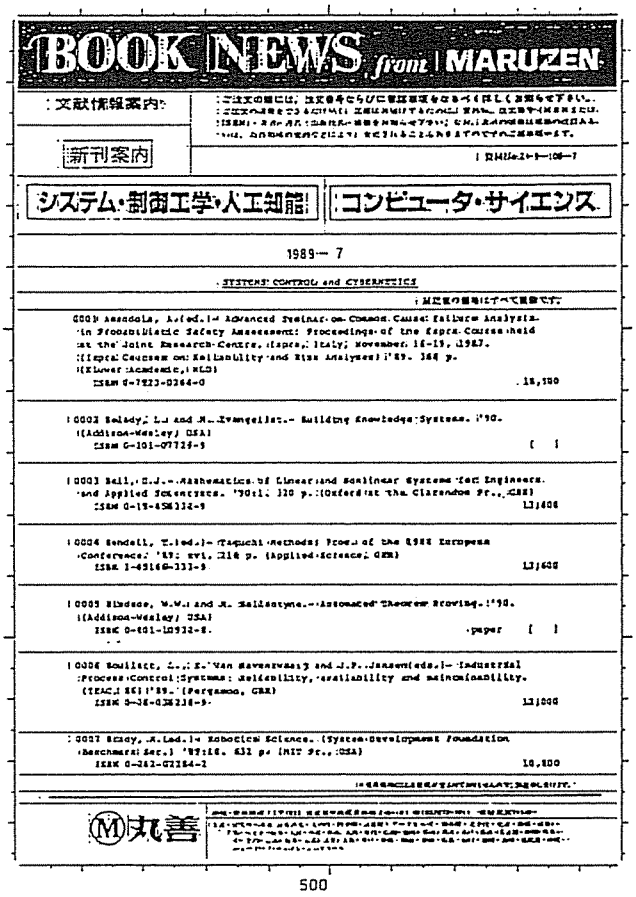


(a) reduced cataloging cards on right-lower corner

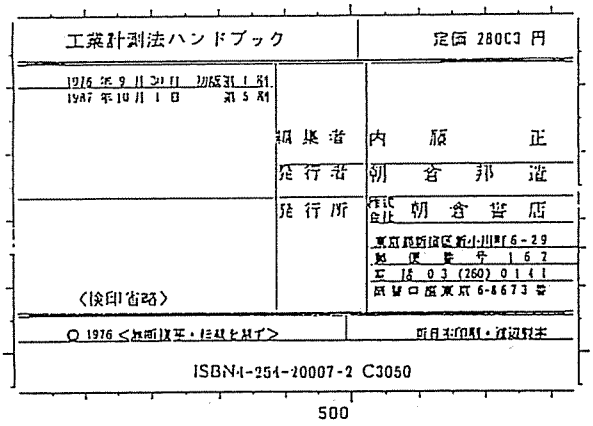


(b) reduced cataloging cards on left-upper corner

Fig.12 Recognition of transformed cataloging cards



(a) bibliography guide document



(b) cataloging-in-publication data in book

Fig.13 Recognition of other documents

by Geometric Aspects", Int'l J.of Pattern Recognition and Artificial Intelligence, Vol.2, No.4, pp.641-655(1988).
10) A.DENGEL: "Automatic Visual Classification of Printed Documents", Proc.of MIV-89, pp.276-281(1989).
11) J.HIGASHINO, H.FUJISAWA, Y.NAKANO & M.EJIRI:"A Knowledge-based Segmentation Method for Document Understanding", Proc.of 8th ICPR, pp.745-748(1986).
12) H.YASHIRO, T.MURAKAMI, Y.SHIMA, Y.NAKANO & H.FUJISAWA:"A New Method of Document Structure Extraction Using Generic Layout Knowledge", Proc.of MIV-89, pp.282-287(1989).