INVITED SURVEY PAPER *Surveys on Image Processing Technologies—Algorithms, Sensors and Applications—*

Document Analysis and Recognition

Toyohide WATANABE†, ***Member***

SUMMARY The subject about document image understanding is to extract and classify individual data meaningfully from paper-based documents. Until today, many methods/approaches have been proposed with regard to recognition of various kinds of documents, various technical problems for extensions of OCR, and requirements for practical usages. Of course, though the technical research issues in the early stage are looked upon as complementary attacks for the traditional OCR which is dependent on character recognition techniques, the application ranges or related issues are widely investigated or should be established progressively. This paper addresses current topics about document image understanding from a technical point of view as a survey.

key words: *document model, top-down, bottom-up, layout structure, logical structure, document types, layout recognition*

1. Introduction

The subject about document analysis and recognition concentrates to analyze the document structure physically or logically and also to distinguish individual items organically or interrelatedly. The techniques in many current researches/developments are to extract and classify meaningful information from paper-based documents automatically so as to complementarily support the traditional OCR techniques [1], [2]: roughly speaking, the former mainly focuses on 2-dimensional/1-dimensional recognition (e.g. for page structures, constructive relationships among items) and the latter does chiefly on 0-dimensional recognition (e.g. for characters, items) [3]–[5]. Of course, the system for document analysis and recognition includes the ability of character recognition to enable to manipulate easily the extracted data for the applicable processings.

In comparison with the traditional OCR, the techniques about document analysis and recognition deal with documents totally in point of constructive relationships among items, composition constraints among neighboring/related items, description rules for connection of items, etc. Until today, many methods/approaches have been already proposed or developed in various types/kinds of documents such as tables, banking-checks, journal pages, newspaper pages, application forms, business cards, official letters, official materials, technical reports [6]–[11], as well as various diagrams, drawings, maps, etc. in the graphics recognition [12], [13], and these systems have been implemented. However, currently usable products are not provided as their own complete systems, but are also available as parts of high functional OCR systems. OCR techniques are desirably applicable to the business work in big companies like banks though the practical customers are too limited. This is because the document form also is specified uniformly in the regularly systemalized process or is required that the work for many documents should be done in terms of the performance. Namely, in current situation it is important for the data management in big companies that many documents with the same form should be manipulated effectively and speedily. This situation is not always desirable to researches/developments of document analysis and recognition. The functionality and effectiveness attended inherently with document analysis and recognition must be appealed to a wide range of end utilizations by means of successful products.

In this paper we focus on the technical topics for document analysis and recognition. First, we discuss the document structure from a viewpoint of logical and physical structures. The documents, at least which have been investigated under the researches/developments of document analysis and recognition, can be evaluated with respect to the logical and physical features of document construction. In general, physical features are called layout structures on the basis of the geometric and spatial relationships among composite elements. Second, we discuss the methods which have been proposed for individual documents until today, in accordance with the features of approaches. The knowledge representation means of document structures are roughly classified into frame-based ones and rule-based ones in the traditionally developed methods/approaches. Also, the analysis and recognition methods based on these representations are procedurely organized by means of the interpretation mechanism of individual knowledge representations.

2. Document Structure

Generally, documents are organized systematically more or less under some configuration rules, description notations and so on: individual components which organize each page or whole pages of documents are as-

Manuscript received January 11, 1999.

†The author is with the Department of Information Engineering, Graduate School of Engineering, Nagoya University, Nagoya-shi, 464–8603 Japan.

sociated mutually with constructive relationships. Usually, the constructive relationships are assigned to make individual components meaningful on 2-dimensional space of page or 3-dimensional space of volume from a viewpoint of logical and physical structures. These constructive relationships are specified as a logical structure for the inherent property of item and a layout structure for the locational property of item.

2.1 Logical Structure and Layout Structure

The logical structure and layout structure are basic views for analyzing document configurations [14], [15]. The logical structure defines the content-based properties of individual document components with respect to their relationships among dependent/independent components: usually, the inclusive relationship is effective. While, the layout structure specifies the geometric/spatial positions of individual document components with respect to their physical page structures. In the layout analysis or structure recognition, the knowledge about layout structure is mainly used because this structure can represent the structural positions of individual components explicitly on 2-dimensional space and it is easy to analyze/recognize document components constructively.

Additionally, the layout structure may be classified more or less into the logical and physical layout structures in accordance with the allocation means of individual document components [16], [17]. In some documents such as tables, individual composite elements are allocated into the predefined fields with respect to the attributes regularly. While, in other documents such as article pages, each composite element is located relatively with respect to the physical features of other composite elements (e.g. the sizes, volumes, etc.). Of course, in case of article page, the approximate structure is physically specified in advance. Namely, the logical and physical layout structures take a role to distinguish the features of individual documents. Figure 1 shows a classification of typical documents with respect to the relationship between the logical and physical layout structures. Of course, every existing document is not always distinguished well under such relationships. In Fig. 1, the documents are classified into 4 types [18].

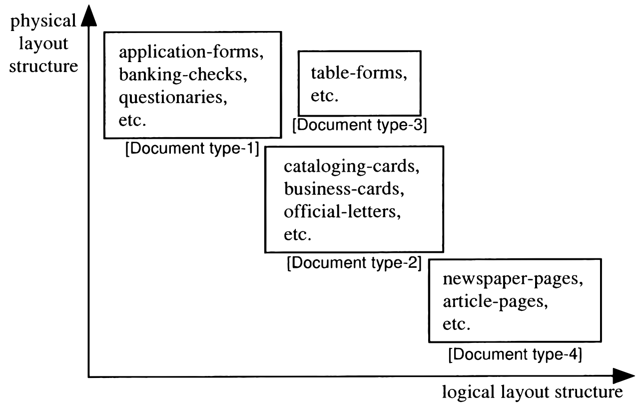


Fig. 1 Document types.

[Document type-1]

This type is too strongly dependent on the physical layout structure. The positions, lengths and so on of individual items are always fixed in advance. For example, application-forms, banking checks, and questionaries are typical.

[Document type-2]

This type is specified by the logical layout structure more effectively than by the physical layout structure. Namely, the position of each item may be moved up/down or left/right from the normal location according to the interrelations among mutually related items or among previously allocated items. For example, cataloging cards, letters, and business cards are typical.

[Document type-3]

This type is dependent on the physical layout structure, as well as the document type-1. In comparison with the document type-1, this type may be complex in the structure (including hierarchical or repeated items) or the layout structure may be guided by other elements (including line segments, blank areas, etc.). The position, length and so on of each item are almost fixed. For example, table-forms are typical.

[Document type-4]

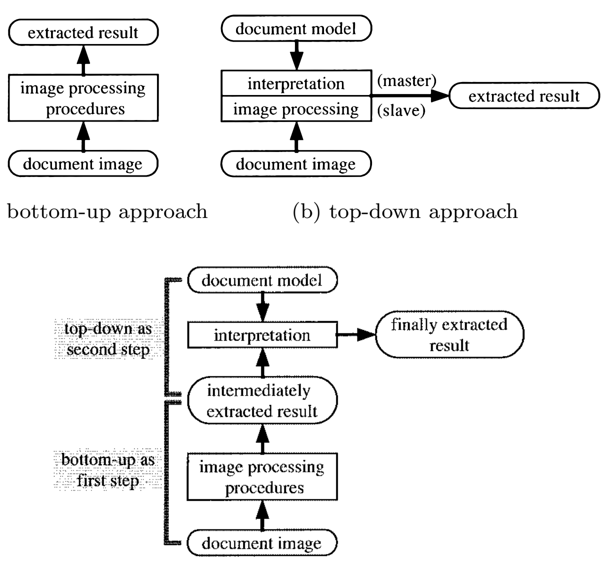
In this type, the positions, lengths and so on of individual items are ordinarily dependent on those of related items or other items. Generally, though the global document structure is predefined by the physical layout structure in this type of document, the allocation strategy for the practical locations of individual items is wholly specified by the constructive relationship among individual items. Namely, this type is related to the physical layout structure in terms of the whole output forms, and also is arranged by the logical layout structure with respect to the locations and shapes of individual items. For example, newspaper-pages and article-pages are typical.

Thus, we observe that different types of documents should be better processed so as to be consistent to the characteristics among items with respect to the logical layout structure and physical layout structure. At least, the optimal recognition methods make sure of the processing efficiency and recognition ratio as well as the flexibility, adaptability and applicability of processing mechanism.

2.2 Approach

Until today, many approaches/methods have been proposed for various kinds of documents. Of course, these approaches/methods were more or less investigated by depending on application-specific document structures. The approach is categorized into the bottom-up and top-down ones [14], [15], [18]. The bottom-up approach has been traditionally applied under the image processing techniques, while the top-down approach was addressed popularly in about 10 years ago under the knowledge-based processing. In the top-down approach, it is important to specify knowledge about document features declaratively, which helps the analysis of structure and the interpretation of individually extracted items. This knowledge is called the document model.

However, it is not easy to define this document model directly for various types of documents or complex documents. Thus, the hybrid approach which is composed complementarily of top-down and bottom-up approaches, is effective in many cases: in the first step the bottom-up approach is applied to the document images in order to approximately analyze/extract the composite elements (as the transformation from pixel-based data to vector/symbolic-based data); and then in the second step the top-down approach is applicable to the interpretation of transformed vector/symbolic-based data with the document model. Of course, as another framework the bottom-up and top-down approaches may be cooperatively integrated. Figure 2 shows such individual approaches conceptually.



(a) bottom-up approach (b) top-down approach

(c) hybrid approach

Fig. 2 Top-down and bottom-up approaches.

2.3 Document Model

The document model, which takes important play in the top-down approach as shown in Fig. 2, specifies the constructive features about document configuration, composition rule, description rule, data domain and so on. Namely, the document model is a kind of knowledge about document structure [14]–[18]. Although it is better to define this document model more generally in order to apply many kinds of documents effectively, the currently proposed frameworks for specifying document models depend on application-specific or similar document classes. This is because it is difficult to interpret document structures successfully by using the more abstracted document model. Additionally, the description information in document model is distinguished from the logical information and physical information [18]–[20].

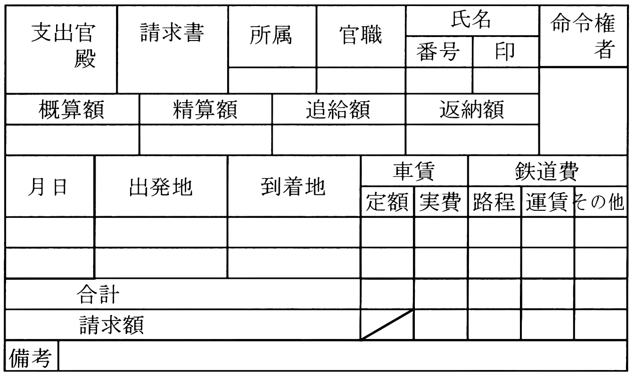
a) Physical or logical representation

The layout structures of documents are dependent on application-specific usages. The difference between the physical representation and logical representation for knowledge specification is dominated with respect to 2-dimensional layout structures of documents. The physical representation is defined by knowledge specification means, which make use of coordinate data of individual items such as positions, sizes, lengths, etc. Logical representation is constructively specified by means of interrelated and interdependent relationships among individual items.

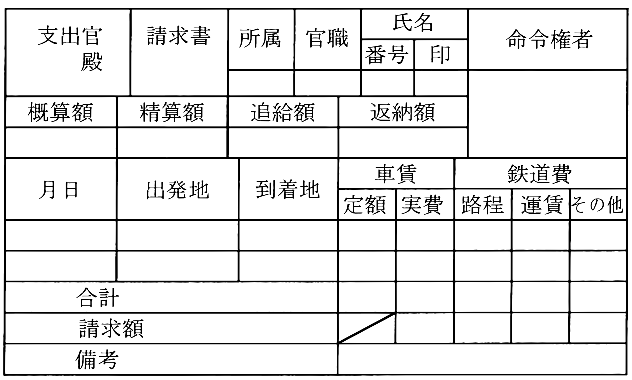
For example, consider two table-forms in Fig. 3. These documents are different in their physical representations but may be the same in their logical representations. Of course, these representations are dependent on various specification views for document structures. However, logical representation is abstracted more than physical representation. The logical representation of document structure is very applicable to various documents of the same or similar types (or classes) if the inference mechanism, based on defined knowledge, is effectively provided. For example, consider two document fragments in Fig. 4. In the logical representation, the predicate "neighbor(A,B)" directly shows that item blocks (or fields) "A" and "B" are adjacent, and is applicable to (a) and (b). In the physical representation, the coordinate data of (a) and (b) must be checked interpretatively whether they are the same positions for x and y-axes. Of course, the meanings of neighboring relationships may be defined in accordance with individual processing and interpretation schema.

b) Abstraction level

In specifying knowledge about documents, we can con-



(a)



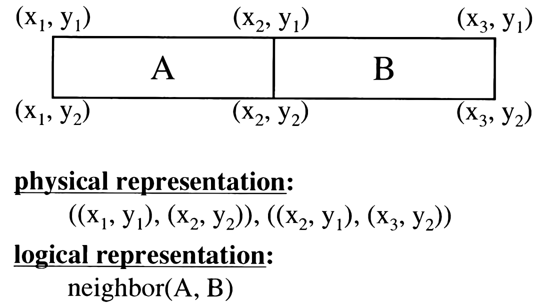
(b)

Fig. 3 Example of table-forms.

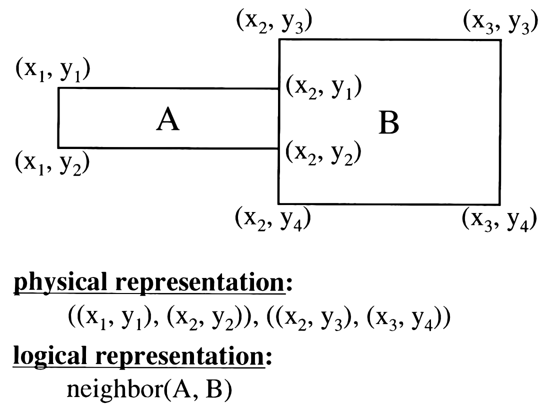
centrate on various characteristics of documents and represent them as usable knowledge [19]–[22]. In this case, the representation method depends on the use of knowledge and the interpretation of document images. We call the representation range of knowledge the abstraction level. The higher the abstraction level is, the stronger the applicability of knowledge becomes. For example, consider the fragments of table-forms shown in Fig. 5.

These fragments are specified as the neighboring relationships among rectangular item blocks (or fields): "A," "B," "C" and "D." In Fig. 5, three different fragment structures are illustrated if we represented the adjacency relationships with commonly shared line segments among two rectangular item blocks (or fields), we can specify the relationships as illustrated in Fig. 6. The arrow indicates the relationship, and the symbols "h" and "v," which are attended with arrows, show whether these neighboring blocks (or fields) are connected horizontally or vertically. Individual document fragments are represented by different neighboring relationships, respectively.

If we make use of the upper-left corners of individual rectangular item blocks (or fields), the resulting adjacency relationships are illustrated in Fig. 7. In this figure, these different document fragments are specified by the same notations as those in Fig. 6. The distinction between the representation in Fig. 6 and that in Fig. 7 is clear. Although the representation for adjacency relationships in Fig. 6 is different in accordance

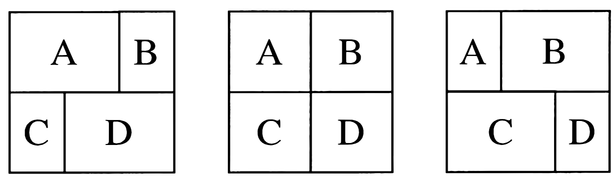


(a)



(b)

Fig. 4 Logical and physical representations.



(a) (b) (c)

Fig. 5 Fragments of table-forms.

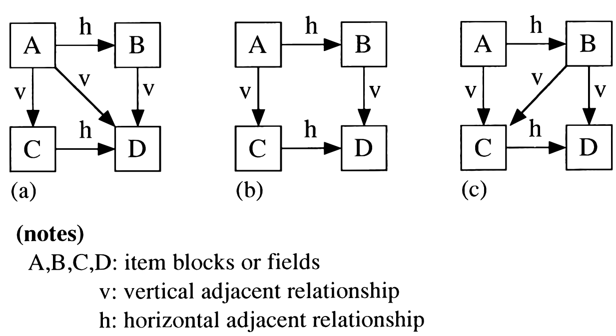


Fig. 6 Representation by commonly shared line segments.

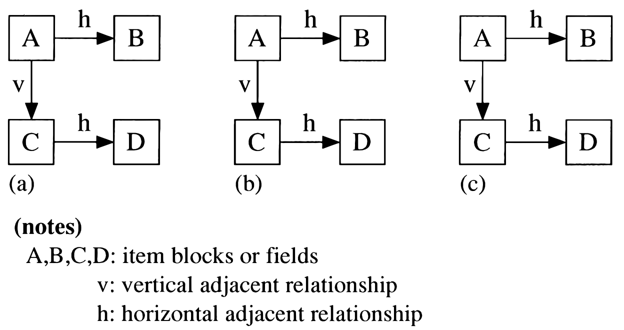


Fig. 7 Representation by upper-left corners.

with their geometric structures, those in Fig. 7 are the same even if their geometric relationships were different. Of course, Fig. 5 (b) may be transformed into another adjacency relationship. Item block "B" is a neighbor of item block "D" by means of the arrow "v" in place of the arrow "h" between item blocks "C' and "D." Such a difference is derived from specification views of individual knowledge designers. Thus, both specifications are true for the layout analyzer: in our case, it is only one processing purpose that the layout analyzer judges whether the currently parsing layout structure is consistent with the predefined knowledge of this table-form. At least, we can conclude that the representation means in Fig. 7 is more strongly abstracted than that in Fig. 6.

c) Semantic or syntactic information

Currently, knowledge which is usable for analyzing and recognizing document images is almost composed on the basis of the layout structures of documents. This is because the currently proposed methods focus mainly on the geometric characteristics of document structures and distinguish individual items with the assistance of structure analysis one by one. However, these methods based on the models of layout structures are not always applicable to various documents: for example, consider the cases where the layout structures are irregularly transformed from the original forms or where the structural characteristics are not extracted sufficiently.

In general, the recognition results, which were distinguished only by means of applications of syntactic knowledge, may not be always correct. This is because the processing based on the syntactic knowledge is too heavily dependent on the representation means of objects. Thus, we must investigate other knowledge with a view to resolving such a problem. We call it semantic knowledge [19], [20]. Semantic knowledge is information which defines the domains of individual items and specifies the interdependent and interrelated relationships among items. Semantic information is complementary to syntactic information with a view to understanding document images. The issue about document image understanding not only focuses on the development of an effective method for extracting and classifying individual items automatically, but also must concentrate on making sure that the identified results are valid.

3. Document Model and Analysis Method

Generally, the recognition of document structure can be looked upon as the paradigm, illustrated in Fig. 8. The document class recognition may be not addressed under the framework of document analysis and recognition explicitly [23]. Many researches/developments excluded this document class recognition as the direct subject because the documents to be analyzed are too strongly dependent on the application and also the main features about layout structures of documents are explicitly prespecified. The layout recognition, item recognition and character recognition are important modules to extract and classify the meaningful information from the document images, though the document class recognition takes an important role to expand to a wide-range of applications progressively [3]–[5], [24].

The layout recognition separates individual item groups under the constructive relationship and is the most critical procedure in the document analysis and recognition. This procedure divides 2-dimensional document data into groups of 1-dimensional item sequence data. The item recognition identifies individual items as meaningful composite elements of documents under the descriptive relationship. This procedure transforms 1-dimensional item sequence data into ordered (or unordered) collections of 0-dimensional character data. Finally, the character recognition extracts individual characters/symbols as meaningful words/notations under the data formats or domain values. Of course, these procedures may be not always organized systematically or the objects to be recognized may be not always determined uniformly. This is because the recognition procedure is dependent on the specification of document model. Some methods do not separate these three layer procedures independently; and other methods organize three layer recognition procedures explicitly. The layout recognition and item recognition are composed under the complementary organization, and the recognition level between these procedures is not explicit: in one method the item recognition is not explicitly provided when the item sequence rule is very simple or when items are easily classified in the layout recognition.

The document model is classified into frame-based means and rule-based means, as the knowledge representation. Also, the frame-based means is divided into list-based means, tree-based means and graph-based means [14], [18]. Generally, the list-based means is adaptable directly to analyze the structure, and also is more effective because the knowledge representation is very simple. On the other hand, the rule-based means is applicable to analyze more complicated docu-

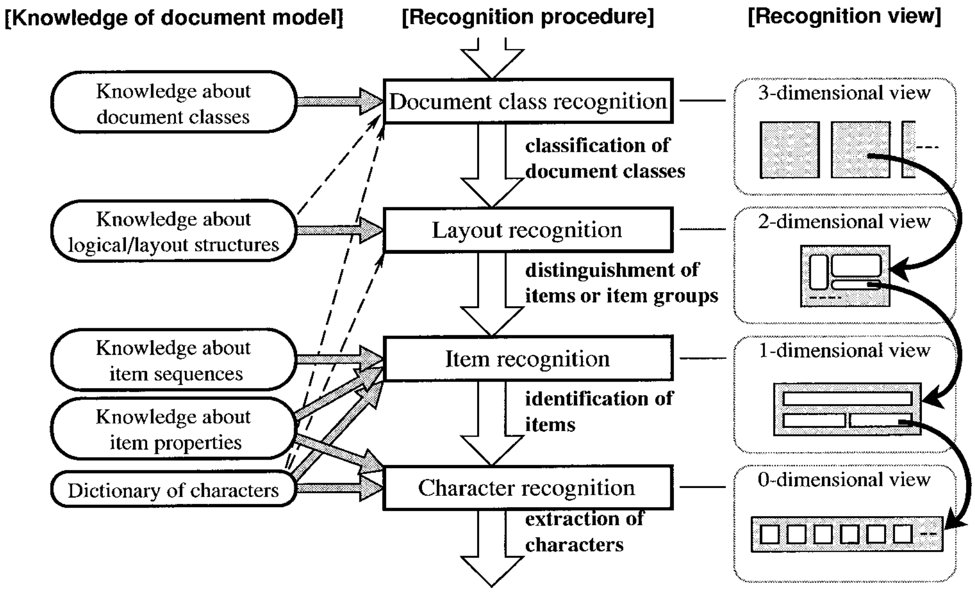


Fig. 8 Framework of document image processing.

ment structure, because individual items as composite elements of documents are inferred through the constructive relationships among individually adjacent items [25]–[27]. Namely, in the document types shown in Fig. 1 the documents formed deterministically by the physical layout structure are possibly specified by means of list-based knowledge representation. On the contrary, the documents specified irregularly by the logical layout structure should be analyzed through rule-based knowledge representation.

Concerning the layout analysis methods, the recognition paradigm is shown with respect to individual document types, hereafter. The layout recognition based on the layout knowledge must analyze/interpret the document structure to identify individual item data with the applicable document model, which specifies not only layout knowledge but also other various information about the documents. Namely, the knowledge-based layout recognition process is organized systematically as the model-driven approach. Here, we address different approaches for various types of documents and arrange individual frameworks from viewpoints of knowledge representation and processing mechanism.

[Document type-1]

All item data are always located to the predefined positions, often associated with some leading words (or key terms) and so on. Basically, individual data items can be assigned line by line. Thus, the layout knowledge is very simply and certainly specified on the basis of the locations of individual data items. The structure of layout knowledge is representable by the list

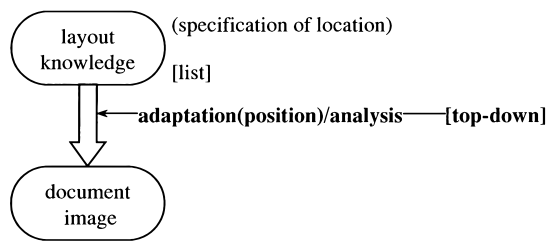


Fig. 9 Framework of document type-1.

(or frame). In this case, the layout recognition process can be organized conceptually as shown in Fig. 9. The document images are analyzed directly by appropriate image processing routines under the adaptation of layout knowledge. Of course, this process is mainly employed in the top-down approach.

For example, the method for the application-forms, proposed in [28], is typical. And, many traditionally proposed knowledge-based (or model-driven) approaches are categorized into the processings for this type of document with the knowledge representation means.

[Document type-2]

In this type of document, individual data items are controlled mainly by the logical layout structure with respect to the location: the positions of data items are variously alterable by other related data items. Thus, the layout knowledge cannot accommodate the locational information effectively. Though some methods which used coordinate data to assign the positions of individual data items had been reported, these meth-

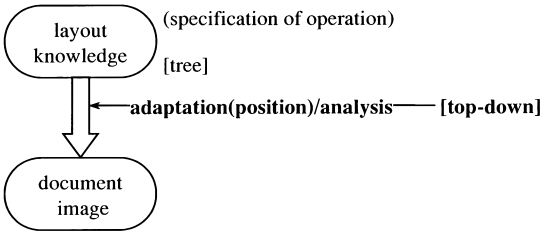


Fig. 10 Framework of document type-2.

ods are too strongly limited concerning the adaptability: of course, these methods introduced the reasonable matching ranges for the preassigned coordinate data in order to get rid of disadvantages for the adaptability, but they were unsuccessful to apply well to variously organized instances. Namely, in comparison with the document type-1, it is better that the layout knowledge in the document type-2 should exclude coordinate data (as physical information) such as the positions, sizes, lengths, etc. [19], [20].

For example, the method, proposed in [4], [24] is very applicable to various documents of this type: business cards, cataloging cards, reference lists, etc. This layout knowledge is defined by the tree as a sequence set of partition operations, based on the neighboring relationships among individual data items (as logical information) [29], though many conventional methods adapt ad hoc means which specify the constructive structure, coordinate data and so on about individual data items (or physical properties of data items) directly. Figure 10 shows the layout recognition process based on such operations, conceptually. Of course, this process is composed in the top-down approach, as well as that of the document type-1.

[Document type-3]

This type of document is commonly organized as the table-form and individual data items may be allocated to the predefined positions as well as the document type-1. Moreover, individual data items may relate to other data items hierarchically or repeatedly [30]. In some cases these data items are surrounded with vertical/horizontal line segments [31] or in other cases they are separated by blanks [32]. Of course, this type of document is constrained strongly by the physical layout structure. The layout knowledge can be represented by the tree because the hierarchical structures among data items can be regarded as the upper-lower relationship among nodes in the tree and the repeating structures can also be defined as the attribute of the upper node [33], [34].

For example, the method, proposed in [19]–[21], [23], [31], [32], specifies the complex structure, using two different binary trees: global structure tree and local structure tree. The global structure tree specifies the characteristic structure such as hierarchy and repeating in addition to the constructive relationships among data items/groups. Also, the local structure tree defines in detail the connectivities among data items individually. Figure 11 shows a layout recognition process conceptually. In this figure, two different processing phases are illustrated: first, the analysis phase extracts the characteristic points (as upper-left corners of data item fields, generated from vertical and horizontal line segments) from document images by means of the bottom-up approach, and then the interpretation phase distinguishes the extracted characteristic points with the layout knowledge in the top-down approach.

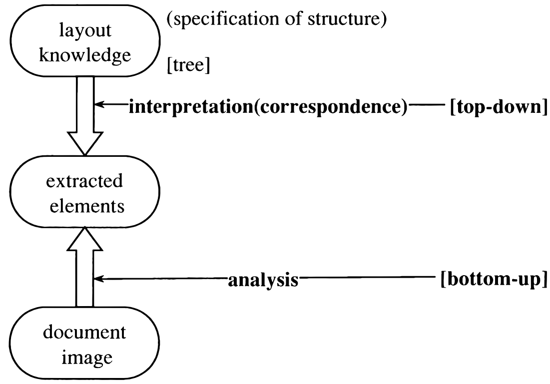


Fig. 11 Framework of document type-3.

[Document type-4]

In this type of document, individual data items and their composite elements are more complicatedly interrelated than the previously addressed document types. The positions of individual data items or composite elements are too strongly dependent on the previously allocated data items or other related composite elements under the physical layout structure. The layout knowledge can be defined by the rules.

For example, the method for the layout recognition of Japanese newspapers, proposed in [35], or the method for the layout recognition of English newspapers, proposed in [36], is composed of production system in order to establish the correct correspondence among interrelated data items/elements. The framework of this layout recognition is shown in Fig. 12. The approximate processing mechanism is the same as that in the document type-3: first, the analysis/labeling phase extracts candidates for composite elements or data items from document images and organizes the extracted candidates as a graph/list to determine the constructive relationship in the bottom-up manner; and second, in the top-down manner the interpretation phase interprets the constructive links among the extracted candidates (as a graph/list) with the layout knowledge (as a set of rules) and distinguishes the structure by making the connectivities among data items and/or composite elements clear. In this case, the layout meta-knowledge was used in [35] with a view to making the control for rule interpretation easy. This is because the layout meta-knowledge can make the interpretation efficiency effective by classifying different rules into the similarly related sets. Of course, this layout meta-knowledge itself is also represented as the rules. The method for the layout recognition of article-pages [37] is simpler than those of these newspapers because the physical layout structure in articles is more explicit than that of newspapers and also the logical layout structure is simpler than that of newspapers.

4. Document Type and Layout Knowledge

In order to apply appropriate representation means of layout knowledge and the processing mechanisms to various types of documents correspondingly, it is better to make the recognition ratio and processing efficiency high. Of course, the recognition process is more or less dependent on the representation means of layout knowledge. In this case, we must pay attention that the representation means of layout knowledge, addressed in Sect. 3, are not always applicable to the corresponding document types, but other means may be usable to some document types. However, it is very important for us to focus on the relationship between document types and knowledge representation means. This is because the more strongly documents are dependent on the logical layout structure, the more complex the knowledge representation means become.

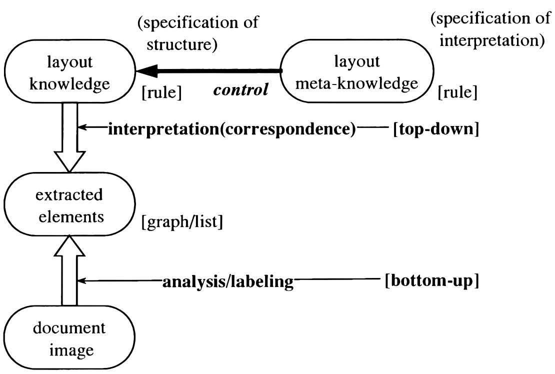


Fig. 12 Framework of document type-4.

Figure 13 shows that individual document types are organized under the relationship between the physical layout structure and logical layout structure approximately. Additionally, in this figure these document types are arranged in accordance with the representation means of layout knowledge: especially, in the document type-3 (such as business cards [4], cataloging cards [30], business/official letters [38], etc.) the layout knowledge can be defined as a set of operations for partitioning regions hierarchically though the layout knowledge in the other document types specifies the properties of individual data items or/and relationships among data items directly. This is because the documents of document type-3 are almost independent of the physical layout structures and are possibly derived from the logical layout structures.

Furthermore, we discuss the relationship between the logical layout structure and physical layout structure from a viewpoint of knowledge representation means. In four types of documents, various kinds of knowledge representation means are used: lists, trees, graphs (or networks) and rules. The list, tree and graph are fundamentally frame-based representation means, and these are smartly applicable when the document structure is explicitly definable under the physical layout structure, except the document type-2 (in the document type-2 the knowledge representation means are not to specify the locations and/or structures, but to specify the operations).

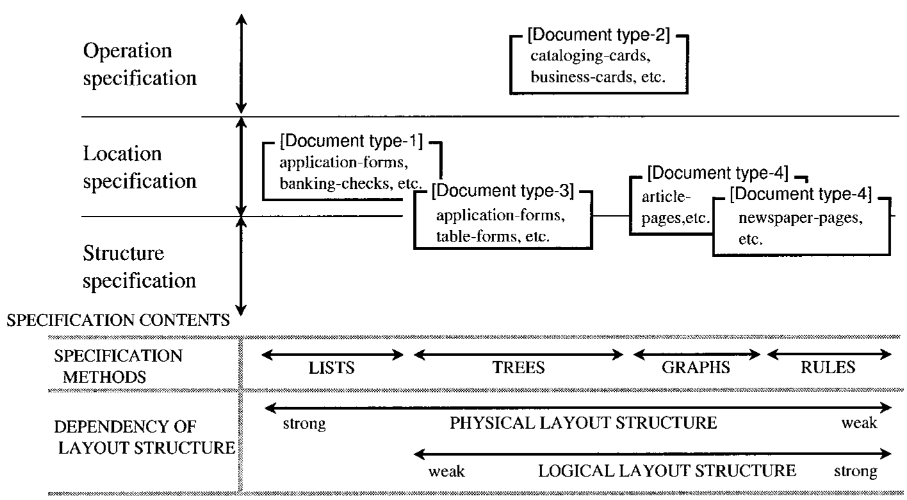


Fig. 13 Representation means of layout structures.

If the document structure could be defined strongly by the physical layout structure, it is better to make use of simpler frame-based representation means: the list is more successful than the tree or graph in the document type-1 because the processing efficiency is superior, and the knowledge definition/management is easy; and the tree is more effective than the graph in the document type-3, similarly. Of course, in case of the document type-3 the graph is adaptable but such selection is not smart. In comparison with these frame-based layout knowledge representation, the rule-based knowledge representation is very powerful. However, this selection often generates drawbacks for the processing efficiency because the knowledge representation becomes complex and also the interpretation phase exhausts much time to navigate on many possibilities.

When we try to analyze/recognize the layout structures of some documents, it is very important to make use of simpler knowledge representation means, as possible as we can, for most basic and characteristic properties of document structures, and then specify other properties effectively which are inherently dependent on individual documents.

5. Other Issues

In this paper, we mainly addressed model-based document analysis and recognition approaches from viewpoints of knowledge representations and analysis methods. However, the research/development issues, attended with document analysis and recognition are not limited on the discussion points in this paper, but extend to various kinds of documents (e.g. maps, diagrams, music books, engineering drawings, business graphs, etc.) [12], [13], [39]–[41], data driven approaches for technical problems (e.g. segmentation of complicated composite document elements, separation among touched characters or line segments, extraction of character strings from attended backgrounds, etc.), analysis/recognition of complex documents or mixed types of documents, recognition of handwritting characters from filled-in types of documents, document processing for applications or on applications (e.g. OCR, information retrieval, WWW server, etc.) and so on [6]–[13].

Of course, these topics related to the subject of document analysis and recognition have to investigate as a fundamental means for the construction and utilization of usable information bases in order to make the future direction clear. Currently, the results derived from researches/developments in this document analysis and recognition are not always usable or applicable as their own complete systems/packages. The document analysis and recognition is looked upon as a functional module of OCR, and takes an important role in the data input tasks.

6. Conclusion

In this paper, we addressed the layout recognition subject. Many documents have their own layout structures more or less, and the layout recognition methods work well, using the layout knowledge which has to be established effectively from such inherent layout structures under the model-driven (or top-down) approach. Such a framework is very smart in point of the flexibility, applicability and adaptability of processing abilities, and also makes the processing efficiency and recognition ratio high logically.

The research/development history for document analysis and recognition counts up about 20 years. Also, many useful products have been proposed: in particular, character recognition and drawing interpretation. Additionally, since 1990 the knowledge-based approach based on the model-based understanding paradigm has mainly been investigated eagerly, and currently presents most basic framework. Of course, the approaches/methods related closely or complementarily to this approach are also investigated in a wide range of applications with a view to analyzing/recognizing various kinds of documents, attaining high recognition ratio, making the processing efficiency successful, and making more complicated documents effectively. At least, as one of the future topics for the document analysis and recognition, it is so desirable that these subjects should be expanded in a wide range of applications.

References

[1] H.S. Baird, "Anatomy of a page reader," Proc. MVA '90, pp.483–486, 1990.
[2] M. Ejiri, "Knowledge-based approaches to practical image processing," Proc. MIV '89, pp.1–8, 1989.
[3] T. Watanabe and T. Fukumura, "An architectural framework in document image understanding," Proc. ICARCV '94, pp.1431–1435, 1994.
[4] J. Higashino, H. Fujisawa, Y. Nakano, and M. Ejiri, "A knowledge-based segmentation method for document understanding," Proc. ICPR '86, pp.745–748, 1986.
[5] T. Watanabe, Q. Luo, Y. Yoshida, and Y. Inagaki, "A stepwise recognition method of library cataloging cards on the basis of various kinds of knowledge," Proc. IPCCC '90, pp.821–827, 1990.
[6] L. O'Gorman and R. Kasturi, "Document image analysis systems," IEEE Comput., vol.25, no.7, pp.5–8, 1992.
[7] L. O'Gorman and R. Kasturi, "Document Image Analysis," IEEE Computer Society Press, 1995.
[8] IAPR TC-10/11, "Proc. of ICDAR '91, vol.1/vol.2," AFCET-IRISA/INRIA, 1991.
[9] IAPR TC-10/11, "Proc. of ICDAR '93," IEEE Computer Society Press, 1993.
[10] IAPR TC-10/11, "Proc. of ICDAR '95, vol.1/vol.2," IEEE Computer Society Press, 1995.
[11] IAPR TC-10/11, "Proc. of ICDAR '97, vol.1/vol.2," IEEE Computer Society Press, 1997.
[12] R. Kasturi and K. Tombre, eds., "Graphics recognition:

Methods and applications," Lecture Notes in Computer Science 1072, 1996.
[13] K. Tombre and A.K. Chhabra, eds., "Graphics recognition: Algorithms and systems," Lecture Notes in Computer Science 1389, 1998.
[14] T. Watanabe, Q. Luo, and N. Sugie, "Structure recognition method for various types of documents," Int'l J. of Machine Vision and Applications, vol.6, pp.163–176, 1993.
[15] T. Watanabe, Q. Luo, and T. Fukumura, "A framework of layout recognition of document understanding," Proc. SDAIR '92, pp.77–95, 1992.
[16] H. Masai and T. Watanabe, "Identification of document types from various kinds of document images based on physical and layout features," Proc. MVA '96, pp.369–372, 1996.
[17] H. Masai and T. Watanabe, "Document categorization for document image understanding," Proc. ACCV '98 (Lecture Notes in Computer Science 1352), pp.105–112, 1998.
[18] T. Watanabe, "A guide-line for specifying layout knowledge," Proc. SPIE/EI '99, pp.162–172, 1999.
[19] T. Watanabe and Q. Luo, "A multi-layer recognition method for understanding table-form documents," Int'l J. Imaging Systems and Technology, vol.7, pp.279–288, 1996.
[20] T. Watanabe and T. Fukumura, "A framework for validating recognized results in understanding table-form document images," Proc. ICDAR '95, pp.536–539, 1995.
[21] T. Watanabe, Q. Luo, and N. Sugie, "Knowledge for understanding table-form documents," IEICE Trans. Inf. & Syst., vol.E77-D, no.7, pp.761–769, July 1994.
[22] T. Watanabe, Q. Luo, and N. Sugie, "Toward a practical document understanding of table-form documents: Its framework and knowledge representation," Proc. ICDAR '93, pp.510–515, 1993.
[23] T. Watanabe, Q. Luo, and N. Sugie, "Layout recognition of multi-kinds of table-form documents," IEEE Trans. Pattern Anal. & Maci. Intell., vol.17, no.4, pp.432–445, 1995.
[24] Q. Luo, T. Watanabe, Y. Yoshida, and Y. Inagaki, "Recognition of document structure on the basis of spatial and geometric relationships between document items," Proc. MVA '90, pp.461–464, 1990.
[25] D. Niyogi and S. Srihari, "A rule-based system for document understanding," Proc. AAAI '86, pp.789–793, 1986.
[26] F. Esposito, D. Malerba, G. Semeraro, E. Annese, and G. Scafaro, "An experimental page layout recognition system for office document automatic classification: An integrated approach for inductive generalization," Proc. ICPR '90, pp.557–562, 1990.
[27] J.L. Fisher, S.C. Hinds, and D.P. D'amatoi, "A rule-based system for document image segmentation," Proc. ICPR '90, pp.567–572, 1990.
[28] J. Higashino, H. Fujisawa, Y. Nakano, and M. Ejiri, "A knowledge-based segmentation method for document understanding," Proc. ICPR '86, pp.745–748, 1986.
[29] G. Nagy, "Hierarchical representation of optical scanned documents," Proc. ICPR '86, pp.347–349, 1986.
[30] C.F. Lin and C.-Y. Hsiao, "Structural recognition for table-form documents using relaxation techniques," Int'l J. Pattern Recognition and Artificial Intelligence, vol.12, no.7, pp.985–1005, 1998.
[31] T. Watanabe, H. Naruse, Q. Luo, and N. Sugie, "Structure analysis of table-form documents on the basis of the recognition of vertical and horizontal line segments," Proc. ICDAR '91, pp.638–646, 1991.
[32] T. Sobue and T. Watanabe, "Identification of item fields in table-form documents with/without line segments," Proc. MVA '96, pp.522–525, 1996.
[33] Q. Luo, T. Watanabe, and N. Sugie, "Structure recognition of table-form documents on the basis of the automatic acquisition of layout knowledge," Proc. MVA '92, pp.79–82, 1992.
[34] H. Kojima and T. Akiyama, "Table recognition for automatic document entry system," Proc. SPIE '90, pp.285–292, 1990.
[35] Q. Luo, T. Watanabe, and N. Sugie, "A structure recognition method for Japanese newspapers," Proc. SDAIR '92, pp.217–234, 1992.
[36] S. Tsujimoto and H. Asada, "Understanding multi-articled documents," Proc. ICPR '90, pp.551–556, 1990.
[37] K. Kise, K. Momota, M. Yamaoka, J. Sugiyama, N. Babaguchi, and Y. Tezuka, "Model based understanding of document images," Proc. MVA '90, pp.471–474, 1990.
[38] A. Dengel and G. Barth, "High level document analysis guided by geometric aspects," Int'l J. Pattern Recognition and Artificial Intelligence, vol.2, no.4, pp.641–655, 1988.
[39] N. Yokokura and T. Watanabe, "Recognition of composite elements in bar graphs," Proc. MVA '96, pp.348–351, 1996.
[40] N. Yokokura and T. Watanabe, "Recognition of various bar-graph stuctures based on layout model," Proc. ACCV '98 (Lecture Notes in Computer Science 1352), pp.113–120, 1998.
[41] N. Yokokura and T. Watanabe, "Layout-based approach for extracting constructive elements of bar-charts," Proc. GREC '97 (Lecture Notes in Computer Science 1389), pp.163–174, 1998.

Toyohide Watanabe received the B.S., M.S., and Ph.D. degrees from Kyoto University in 1972, 1974, and 1983, respectively. In 1987, he became an associate professor in Department of Information Engineering at Nagoya University. He is currently a professor in Department of Information Engineering of Graduate School of Engineering at Nagoya University. Dr. Watanabe's research interests include the knowledge/data engineering, parallel and distributed process interaction, and document image understanding and map/drawing interpretation. He is a member of the Information Processing Society of Japan, Institute of Electronics, Information, and Communication Engineers of Japan, Japan Society for Software Science, Japan Society of Artificial Intelligence, ACM, AAAI, and IEEE Computer Society.