Mathematical Research

Computer Analysis of Images and Patterns

edited by

K. Voss
D. Chetverikov
G. Sommer

/olume 55

AKADEMIE-VERLAG BERLIN

Herausgeber:

Prof. Dr. Klaus Voss
Friedrich-Schiller-Universität Jena
Sektion Technologie

Dr. Dmitry Chetverikov
Hungarian Academy of Sciences, Budapest
Computer and Automation Institute

Dr. Gerald Sommer
Friedrich-Schiller-Universität Jena
Sektion Technologie

Die Titel dieser Schriftenreihe werden vom Originalmanuskript der Autoren reproduziert.

ISBN 3-05-500755-7
ISSN 0138-3019

Erschienen im Akademie-Verlag Berlin, Leipziger Str. 3-4, Berlin, DDR-1086

Lizenznummer: 202·100/415/89
Printed in the German Democratic Republic
Gesamtherstellung: VEB Kongreß- und Werbedruck, Oberlungwitz, DDR-9273
Lektor: Dipl.-Phys. Gisela Lagowitz
LSV 1085
Bestellnummer: 764 071 8 (2182/55)
03500

IMAGE DATABASE MANAGEMENT SYSTEM. APPLIED THEORIES, TOOLS AND DECISIONS

P. Stanchev[1] – F.Rabitti [2]

In this paper we address the problem of building a database system for images (IM_DBMS), in which the images can be accessed by a partial description of the image content. The approach allows a limited automatic analysis, using ARG techniques and a rule-based system for images belonging to a domain described in advance. The semantic objects, recognized in the images during the analysis process, are interpreted according to the Theory of evidence. The image query processing is based on special access structures generated from the image analysis process.

1. INTRODUCTION

When large amounts of images have to be managed in a computer system, the need to apply the database technology naturally arises. In the last decade, much of the work in the field of "image databases" appeared in the proceedings of the "IEEE Workshops on Pictorial Data Description and Management", starting from 1977, and in some other series such as "Computer Graphics and Image Processing", "Computer Vision, Graphics, and Image Processing", "Image and Vision Computing", etc. Some essential sources are: the work edited by Chang and Kunii [7] and the collections of papers on pictorial applications and information systems in [5], [6], [10], [12], etc. The valuable survey of Tamura and Yokoya [20] includes insights into many actual approaches, as well as descriptions of several systems, such as the Graphics-oriented Relational Algebraic Interpreter (GRAIN), the Relational Database system for Images (REDI), the Database system of Microscopic Cell Images (IDB), etc. Another work [8] presents a survey of seven commercial systems, currently available and their software capabilities: Xerox's 8010 Information System, TERA's Automated Records Management Systems, Teknotron's Systems, Scitex's Response 250, Toshiba's Document Image Filing System DF2100, CCA's Spatial Data Management System, IIS's System 600 series of software products.

But, most of the systems, appearing under the heading of image databases are often image systems without full database functionalities or database systems not directly dealing with images [13]. In fact, most of the existing systems are application specific, that is, the way in which images are stored, organized and retrieved is specific of a certain application and cannot be generalized to different applications.

The main conceptual problem in dealing with images derives from the difficulty to exactly define and interpret the content of images. Images can be very rich in semantics, but are subjected to different interpretations according to the human perception of the application domain. On the one hand, it is difficult to recognize the objects (with the associated interpretation) contained in an image, on the other hand is difficult to determine and represent the mutual relationships among these objects, since they form structures which vary greatly from image to image.

For the problem of image retrieval by content, one could think to apply Data Base Management System (DBMS) or Information Retrieval System (IRS) techniques. However, with respect to DBMS's, it is difficult to recognize regular structures of objects contained in images, and then organize image instances into a limited number of types, to which the interpretation is associated. This is the approach required by the strictly typed data models adopted in database systems [21]. In IRS, instead, a free formatting of text is allowed, usually respecting some loose hierarchical structuring in sections, subsections, paragraphs and sentences. These systems do not attempt to understand the text (unless some expert system approach is adopted), but still allow an effective retrieval on text. In fact, as opposed to image objects, they can exactly recognize words (as ASCII patterns), on which they base their retrieval capabilities with the possible help of a thesaurus to support synonyms [18]. This is possible, in case of text, because a common semantic is associated to the words used in the natural language. Hence, both DBMS and IRS approaches cannot be directly applied to image retrieval.

[1] Institute of Mathematics, Bulgarian Academy of Sciences, G. Bonchev Str., 8 - 1113 Sofia (Bulgaria)

[2] Istituto di Elaborazione della Informazione, National Research Council, Via S. Maria, 46 - 56100 Pisa (Italy)

In addressing the problem of image retrieval of stored images, if we want to think of a system trying to do for images what DBMS and IRS do for formatted data and text, we must accept some indeterminantness, characteristic of images, and then deal with the inaccuracy introduced by this fact. In [14] an approach based on Fuzzy set theory has been applied to the analysis and description of pictorial images. Certainty factors for the recognition of objects inside the images are computed using fuzzy logic rules.

A major practical problem for pictorial images is element recognition. The first step in the image analysis and recognition process is the decomposition of the images into relevant and identifiable elements, which are the basic components which will constitute the building blocks of the image structure.

In this step, often called segmentation in pictorial image processing [1], the image space is partitioned into meaningful regions, corresponding to image elements. After the segmentation, the system must recognize the tentative elements in the image, matching them with the pictorial representation of the elements to be searched. In this process, different variations as changes in size, rotation, translation (eg. using discrete Fourier transforms), can be attempted. An additional problem arises when partial element overlapping occurs in the image.

The result of this process should be the set of basic elements recognized and their relative positions. However, this is a highly computing intensive process which often requires special hardware, such as array processors, exploiting the inherent parallelism of the algorithms in order to have acceptable response times. In the end, the system might even not be able to exactly identify the single elements.

In this paper, a database system for images, IM_DBMS, is presented. It supports the analysis and retrieval of images. It is required that the images belong to a specific domain which must be described in advance to the system. A limited automatic analysis of the images is performed before storing the images in the database. This process is accomplished using a rule-based system. The interpretation of the content of the images is based on the Dempster-Shafer theory of evidence [11].

2. THE IM_DBMS SYSTEM

The main data "objects" that are input, analysed, stored, retrieved in IM_DBMS are bit-map images, obtained through image scanning process or graphical images, obtained by the tools of a graphical editor.

Figure 1. shows the architecture for the IM_DBMS. In the system we can identify six interrelated phases: domain description, image input and editing, image analysis, image storage, image retrieval, and image output and communication.

2.1. Domain Description in IM_DBMS

The purpose of this phase is to describe the characteristics of the application domain of the images to be stored and retrieved. The domain description function supplies with initial information the various phases of the image creation and analysis. It comprises:

a) establishment of the domain basic elements, for the puproses of image elements understanding. Domain basic elements are the basic components which will constitute the building blocks of the image structures. Image elements are domain basic elements established during the segmentation and recognized during image element understanding phase.

b) describing the possible relations and attributes of the basic elements, for the purposes of object recognition. Objects are meaningful for the application domain basic elements constructions;

c) indication of representative images, for the purposes of the image interpretation and clustering.

2.2. Image Input and Editing

The puprose of this phase is to enter the image into the computer storage. Bit-map images are

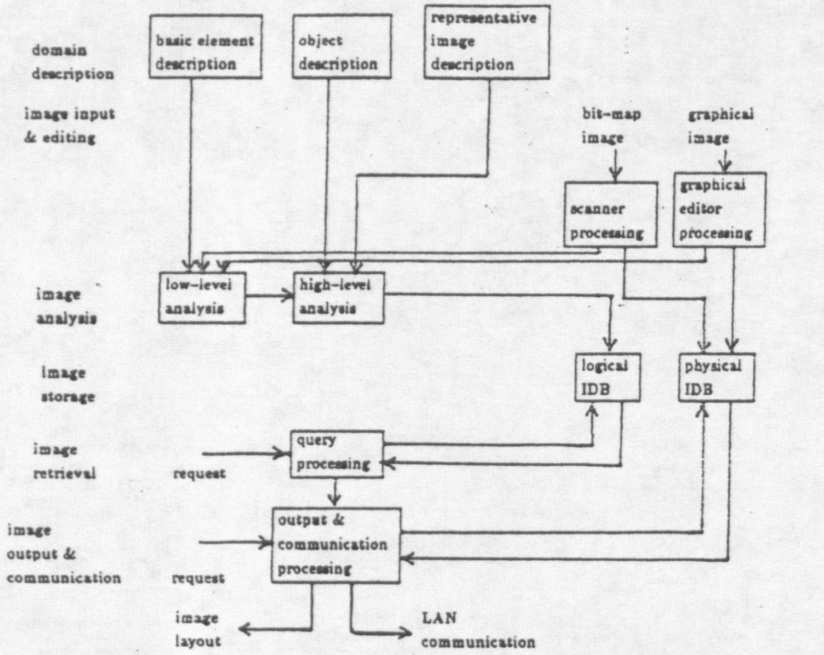

Figure 1. *Architecture for IM_DBMS*

acquired by a scanner and are transformed in image binary array form. Graphical images are created by a graphical editor and are transformed in a set of graphical editor primitives.

2.3. Image Analysis

The process of image analysis is concerned with the global analysis of the images. It could be considered separated into two phases.

The first phase, usually called low-level image analysis, is concerned with the extraction of an adequate form of knowledge from the image data and representing it in some appropriate formalism.

In the second phase, called high-level image analysis, the actual, in the sense of image meaning comprehension, analysis of the image is performed, using the extracted representation. The global image representation is obtained.

2.3.1. Low-Level Image Analysis

The low-level image analysis accepts an image (bit-map or graphical) and recognizes the basic elements in it, their relative positions and the associated distinguishing attributes.

The low-level image analysis is constituted by the following steps:

a) image elements understanding

Since the number of the basic elements (eg. polilines, curves, etc.) can be very large in a single image (in the order of thousands) a very efficient approach is required for recognizing the basic elements which are meaningful in the application domain. They constitute the basic symbols which compose the semantic objects in the image. In this phase, it is not possible to adopt a rule system, based on a generalized inference mechanism with back-tracking, because of its computational complexity. We need instead more efficient and specialized algorithms (with polynomial computation complexity) even if we have to pay this with a description system less rich in semantic content. For this reason we have adopted an approach based on the Attributed Relational Graphs [9].

The ARG graph is a relational structure which consists of a set of nodes and a set of branches representing the relations between the nodes, as both nodes and branches may have some attributes assigned to them.

During the Image Element Analysis, an ARG representation of the image (in terms of basic elements of the application domain, thier relationships and attributes) is obtained and stored in the image description file.

2.3.2. High-Level Image Analysis

The high-level image analysis uses the information about the elements, contained in the image and their relative position, obtained in the low-level image analysis. It is constituted by the following steps:

a) object recognition

The result of this phase is a set of objects recognized in the image.

The purpose of this phase is to recognize more semantically meaningful objects from the basic elements derived and organized in the previous phase. This task is accomplished by recursively applying the production rules defined for the chosen application domain. An inference mechanism based on backward chaining tries to derive from the basic elements more general objects and to give a recognition degree to the object recognized. In this phase a generalized inference mechanism is used. Its computational complexity is acceptable at this point, since fewer objects (in the order of hundreds) are present in the image.

The inference process starts from production rules obtained from the ARG image representation. After this step, a sequence in the form (1) is obtained:

(1)
$$\{O_{1\,1}(\mu_{1\,1}, l_{1\,1}), \ldots, O_{1\,s_1}(\mu_{1\,s_1}, l_{1\,s_1})\}, \ldots,$$
$$\{O_{n\,1}(\mu_{n\,1}, l_{n\,1}), \ldots, O_{n\,s_n}(\mu_{n\,s_n}, l_{n\,s_n})\}$$

Such a sequence describes an image with n distinct physical objects. The unit $O_{i\,j}(\mu_{i\,j}, l_{i\,j})$ is a semantical representation of the physical object $i(i = 1, 2, \ldots, n)$ in the image in the j-th ($j = 1, 2, \ldots, s_j$) recognition (i.e. a semantic object). $\mu_{i\,j}$ and $l_{i\,j}$ are respectively, the recognition degree (RD) and the list of attributes of the i-th physical object in the j-th recognition.

b) image interpretation.

The image interpretation is constituted by a sequence of all interpretations of the objects found in the image.

Using a procedure similar to Barnett's scheme [2], based on the Dempster-Shafer theory of evidence [11], and fully described in [15], we convert the results obtained from the previous phase into a list of new structures, containing information for each object:

(2)
$$\Big\{O_{1\,1}([Bel(O_{1\,1}), 1 - Bel(\overline{O}_{1\,1})], l_{1\,1}), \ldots,$$
$$O_{1\,q_1}([Bel(O_{1\,q_1}), 1 - Bel(\overline{O}_{1\,q_1})], l_{1\,q_1})\Big\}, \ldots,$$
$$\Big\{O_{n\,1}([Bel(O_{n\,1}), 1 - Bel(\overline{O}_{n\,1})], l_{n\,1}), \ldots,$$
$$O_{n\,q_n}([Bel(O_{n\,q_n}), 1 - Bel(\overline{O}_{n\,q_n})], l_{n\,q_n})\Big\}$$

Here $q_i \leq s_i (i = 1,2,\ldots,n)$. The function $Bel(O_{i\ j})$ $(i = 1,2,\ldots,n,\ j = 1,2,\ldots,q_i)$ is a belief function.

The belief function $Bel(O_{i\ j})$ gives the total amount of belief committed to the object $O_{i\ j}$ after all evidence bearing on $O_{i\ j}$ has been pooled. The function Bel provides additional information about $O_{i\ j}$, namely $Bel(\overline{O}_{i\ j})$, the extent to which the evidence supports the negation of $O_{i\ j}$, i.e. $\overline{O}_{i\ j}$. The quantity $1 - Bel(\overline{O}_{i\ j})$ expresses the plausibility of $O_{i\ j}$, i.e., the extent to which the evidence allows one to fail to doubt $O_{i\ j}$. The interval $[Bel(O_{i\ j}), 1 - Bel(\overline{O}_{i\ j})]$ is called belief interval.

In the expression (2), object interpretations with "low" belief (e.g., in the sense of interval mean value less than a chosen one) could be omitted.

c) **image clustering**

The image clustering process is, in principle, similar to the document, clustering of text documents used in Information Retrieval Systems [18]. The most significant classes of images in the application domain are defined in terms of representative images, one for each class. The image interpretations are clustered by comparing them with the class representative images. After this computation the clustering description of the image is expressed as a sequence:

(3)
$$\mu_1, \mu_2, \ldots, \mu_p,$$

where μ_i is the membership degree of the image to the i-th class. .

2.4. Image Storage

The information obtained in the image analysis phase can be used to generate access structures on image content which can be used for efficient image retrieval. Access structures are mainly indices on the objects, contained in all the images, with the associated attributes, and clusters about the interpretations of the images. The images are stored in a physical image data base (physical IDB) and their descriptions in a logical image data base (logical IDB).

The image representation is stored in a file containing, either the graphical image, as sequence of the graphical primitives applied for its composition by the graphical editor, or the binary array representation, in the case of bit-map images obtaining.

The derived image information, resulting from the analysis phase (expressed in terms of the probabilistic model, as composition of objects at different level of complexity, with the associated interval of belief) is stored in an "image header", associated to the image file.

Access structures (that is, the image indices) are built for a fast access to image headers. Two type of indices are constructed:

* Object index. For each object a list OBJ_IND is maintained constituted by a: DOC_ID, OBJ_NAME, OBJ_NUM, ATT, VAL, OBJ_CEN, OBJ_CON, OBJ_OVER, BI1, BI2 where: DOC_ID is a pointer to the image identifier; OBJ_NAME is the object name; OBJ_NUM is the number of the object in the image; ATT and VAL are two arrays, with attribute names and attribute values: OBJ_CEN contains the coordinates of the object center; OBJ_CON is an array with the numbers of the contained objects; OBJ_OVER is an array with the numbers of the overlapping objects: BI1 and BI2 are the bounds of the associated belief interval.
* Cluster index. For each class, defined in the application, a list CLU_IND of elements (DOC_ID, CLU_NAME, DEGREE) is maintained, where DOC_ID is a pointer to an image identifier, corresponding to an image with a non-null degree of membership to this cluster, CLU_NAME is the cluster name, DEGREE is the value of the membership degree.

2.5. Image Retrieval

The developed in IM_DBMS language is a newly developed language for image retrieval by content.

According to the query language, the user specifies the image condition in the form: RETRIEVE IMAGES <image_clause>.

The <image_clause> contains a <cluster_clause> and/or a <object_clause>.

The <cluster_clause> is an "OR" combination of <cluster_predicate>s, each of the form: <cluster_name> <cluster_degree> . This predicate indicates that the images in the database with a similarity to the named class higher than the <cluster_degree> should be retrieved (as requested in the "OR" expression). The <cluster_clause> may be missing if the <object_clause> is present.

The <object_clause> is a boolean combination of <object_group>s. The <object_group> is a group of <object_predicate>s and a <position_relation>. The <object_predicate> has the form: <object_name> <degree>. The <position_relation>s are E, W, N, S, NW, SW, NE, SE, CONTAINING, OVERLAPPING, where E, W, N, S denote the four cardinal points. The object group is interpreted from left to right. The position relation considers only the two objects which it connects. It is fulfilled if it stands for at least one couple of objects in the image.

The <object_clause> must be evaluated, according to he boolean expression, taking into account only the images in the image database, containing those objects, named as <object_name>, with < degree > value, falling into the belief interval. The WITH operator in an <object_predicate> serves the purpose of adding conditions to the attributes associated to the object named as <object_name>. The attribute predicate is a boolean combination of attribute names. In the attribute name comparison with a numerical value, by the operation <, <=, >, >=, = and with a string value, by MATCH operation is allowed.

All the stored images fulfilling the query constitute the query answer set. With this approach the query answers can be ordered by decreasing similarity to the query specification, so an user can receive a ranked output of the retrieved images. (These advantages are typical of the information retrieval techniques).

The formal description of the proposed language and the image query processing algorithm are given in [17].

2.6. Image Output and Communication through Images

After finding the images that satisfy the query specification, corresponding tools for their output (in different colours, shapes, positions, etc.) are required.

The purpose of the communication is to organise the image exchange through local area network.

3. CONCLUSION AND FUTURE WORK

IM_DBMS is under development on SUN/3 workstations, under Unix 4.2. It is being written in C and Quintus Prolog and uses the SUNCORE graphical package [19].

The design of IM_DBMS is based on the experiences of previuos prototypes. The first prototype, described in [15], was based on fuzzy set techniques. It was implemented on IBM PC/AT computer and was intended for the management of business graphical images, generated by a commercial business graphical editor (IBM Graph Assistant). Another prototype, working with graphical images, is described in [16] and is applied for house furnishing design drafts.

We plan to apply the results obtained with IM_DBMS in the project MULTOS, which is part of the ESPRIT (European Strategic PRogramme for Information Technology) [3] and in the project Multimedia information processing supported by the Bulgarian Ministery of culture, science and education. In MULTOS, a first prototype for the storage (based on optical media) and retrieval of multimedia office documents has already been implemented. However, in this prototype images are treated as passive components in the multimedia documents, that is, components which are retrieved as part of the document but cannot actively contribute in the retrieval process (no condition on images can be part of the query, only conditions on attributes, text and the document structure) [4]. In the second MULTOS prototype, we plan to build a specialized subsystem, functionally similar to IM_DBMS, which will allow a higher integration of images in the document retrieval process.

REFERENCES

[1] Ballard, D., and C. Brown: Computer Vision. Prentice Hall 1982.

[2] Barnett, J.: Computational Methods for a Mathematical Theory of Evidence. *Proc. 7-th Inter. Joint Conf. on Artificial Intelligence.* Vancouver, B.C., (1981) 868-875.

[3] Bertino, E., S. Gibbs, F. Rabitti, C. Thanos, and D. Tsichritzis: Architecture of a Multimedia Document Server. *Proc. 2nd ESPRIT Technical Week.* Brussels (1985).

[4] Bertino, E., F. Rabitti, and S. Gibbs: Query Processing in a Multimedia Document System. *ACM Trans. on Office Information Systems.* Vol. 6 (1988) 1-41.

[5] Blaser, A., (Ed.): Database Techniques for Pictorial Applications. Springer Verlag, Berlin 1980.

[6] Chang, S., and K. Fu (Ed.): Pictorial Information Systems. Springer Verlag 1980.

[7] Chang, S., and T. Kunii: Pictorial Data-Base Systems. *IEEE Computer.* Vol.14, 1981.

[8] Chang, S.: Image Information Systems. *Proc. of the IEEE.* Vol. 73, No. 4, (Apr. 1985) 754-764.

[9] Eshera, M., and K. Fu: An Image Understanding Systems Using Attributed Symbolic Representation and Inexact Graph-Matching. *IEEE Trans. Pattern Anal. Macth. Intell.*. Vol. PAMI-8, No.5, (Sept. 1986) 604-618.

[10] Goos, G., and J. Hartmanis (Ed.): Pictorial Information System. Springer Verlag 1980.

[11] Gordon, J., and E. Shortliffe: The Dempster-Shafer Theory of Evidence in Rule-Based Expert Systems. In: Buchanan, B., and E. Shortliffe, (Ed.): Mycin Experiments of the Stanford Heuristic Programming Project. Addison - Wesley Publishing Company (1984) 272-292.

[12] Iyengar, S., and K. Kashyap: Guest Editors' Introduction. Image Databases. *IEEE Transaction on Software Engineering.* Vol. 14, No. 5, (May 1988).

[13] Nagy, G.: Image Database. *Image and Vision Computing.* Vol. 3, No. 3, (Aug. 1985) 111-117.

[14] Rabitti, F., and P. Stanchev: An Approach to Image Retrieval from Large Image Databases. *Proc. ACM-SIGIR Conf.*. New Orleans (1987) 284-295.

[15] Rabitti, F., and P. Stanchev: Graphical Image Retrieval from Large Image Databases. *Proc. AICA Conf.*. Vol. 2, Trento (1987) 69-89.

[16] Rabitti, F., and P. Stanchev: GRIM_DBMS: a Graphical Image DataBase Management System. *Proc. IPIP Working Conf. on Visual Database Systems.* Japan (1989).

[17] Rabitti, F., and P. Stanchev: IM_DBMS: Image Query Language and Image Query Processing. to be published (1989).

[18] Salton, G., and M. McGill: Introduction to Modern Information Retrieval. McGraw-Hill 1983.

[19] SunCore graphical package. *SUN 3 Reference Manuale* 1986.

[20] Tamura, H., and N. Yokoya: Image Database Systems: A Survey. *Pattern Recognition.* Vol.17, No.1 (1984) 29-43.

[21] Tsichritzis, D., and F. Lochovsky: Data Models. Prentice-Hall, Englewood Cliffs, N.J. 1982.

(69.) * Stanchev P., Rabitti F., "Image DataBase Management System. Applied Theories, Tools and Decisions", ***Mathematical Research***, Vol. 55, Academe Verlag, Berlin 1989, 208-214.