Indexing of Images by Pictorial Information

T. Gevers and A.W.M. Smeulders
Faculty of Mathematics & Computer Science, University of Amsterdam
Kruislaan 403, 1098 SJ Amsterdam, The Netherlands
E-mail: gevers@fwi.uva.nl

Abstract

This paper presents the architecture of a system that retrieves images on the basis of pictorial information as opposed to systems based on textual and numerical retrieval.
The pictorial approach is motivated from the observation that the majority of the pictorial information in an image cannot be fully described by textual or numerical information due to its essential limitations in expressive power. Therefore, we propose that the search is carried out from a pictorial specification (i.e. a set of pictorial details and their relationships) where the correspondence is established between the pictorial specification and the images in the database. The method is based on fuzzy Mathematical Morphology operations.
In this paper we concentrate on MRI images of the chest. Experiments have been carried out on a database of those images show encouraging results especially when realizing that the content of medical images is a weakly formalizable domain.

keywords: image retrieval, medical images, pictorial specification, fuzzy Mathematical Morphology.

1 Introduction

Over the past few years, a substantial development is going on in the field of managing non-textual data (e.g. speech and image data) mainly due to the interest of building multimedia information systems (database community) and image database systems (computer vision community) [4], [6]. Managing non-textual data in the form of images include storage, retrieval and processing of pictorial entities. This paper is focussed on the subject of the retrieval of pictorial entities. Image database designs are discussed by Tamura and Yokoya [8] and a survey of commercially available database systems is given by Chang [3].
In this paper an image retrieval system is presented, called Σnigma, which is focussed on searching for details and their relationships on a database containing images as pictorial data. Σnigma retrieves images on the basis of pictorial information as opposed to systems that first reduce the content of the images to textual or numerical abstracts on which the search is defined (i.e. taking the analogy on searching text and numerical databases). We consider images coming from formalizable domains such as electronic schemas and geographical maps as well as from weakly formalizable domains (e.g. medical images).
In this paper, we first present the motivation of the pictorial based retrieval system in Section 2. In Section 3 the architecture of the system is described and in Section 4 experiments carried out on a set of chest images are discussed. Finally, conclusions will be drawn.

2 Motivation of a Pictorial Based Retrieval System

In this paper, image retrieval is defined as a search request on a database of images. The images in the database are "raw" in the sense that no other information is provided except the image as pictorial data. This is the case when no human interaction has been performed to describe the image content when they are entered into the database.
For such a case, to actually do retrieving, the major tasks are the specification of the search request

and the establishing of the correspondence of the search request with all images. One way to approach these tasks is to first reduce the content of all images in the image database as well as the content of the search request to textual and numerical abstracts prior to the search. Once the abstracts have been generated, retrieval is defined as searching the abstracts of the images for correspondence with the search request similar to searching text databases.
Unfortunately, a textual-numerical based approach alone might not be sufficient. This is motivated by the following observations:

- Consider the case that the content of the images in the database is not described by abstracts for some reason.
  Then, a pictorial approach is obvious.
- Even if the content of the images is described and represented by abstracts, there may still be pictorial information that cannot be described by these abstracts due to the essential limitations of text and numbers in expressive power. Take for example, the flow of a river to be described in full detail by text and numbers.
- Also in the case that the image content is described by abstracts, the majority of pictorial information might not be represented by these abstracts due to the semantic richness of pictorial information in images (e.g. to give a complete representation of a house). As a consequence, the textual component might only support a restricted number of queries.
- It is difficult to consistently define and interpret the content of an image, because different abstracts might be generated for the same image when the image is described by different persons.

In this paper, we propose an alternative approach in which the search request is expressed by a pictorial specification and retrieving is defined as the matching of this pictorial specification on the database of images. The correspondence of the pictorial specification and each of the images in the database is expressed by fuzzy Mathematical Morphology [2], [7], which will be discussed in more detail later on.

3 The Architecture of the Image Retrieval System

The architecture of the system is shown in fig. 1.

Figure 1: The architecture of the image retrieval system.

The system contains four databases and three processing units. The four databases are the EIDB (Example Image DataBase), IDB (Image DataBase), LDB (Logical DataBase) and the SDB (Specification DataBase).
The example image database provides example images to the search specification unit allowing extraction of prominent instances in those images. From those instances cartoons are generated. They will be used in the pictorial specification.
The specification database serves as a notebook to hold previously generated pictorial specifications. They may be helpful in more quickly composing new search specifications. In the image database, the images are assembled on which the search is done. They are stored in the database as raw images; no interaction has been performed on them to describe the image content when they are entered into the database.
The logical database, finally, contains the results of previously performed search requests. It holds information about the quality and location of the correspondence between pictorial specifications stored in the specification database and the images in the image database.
The three processing units are the SRSU (Search Request Specification Unit), UI (Indexing Unit) and the EU (Evaluation Unit). In the search request specification unit, a pictorial specification is generated interactively. To be precise, a pictorial specification is made from details and their relationships. A detail is a sign represented by a small image patch. The indexing unit searches for pictorial details in the images of the image database on the basis of the pictorial specification generated in the search request specification unit. In principle, it moves the details over all locations in each image searching for the best fit which does not violate the relationships between the details and stores information about the best fit in the logical database. Finally, retrieval results are interactively evaluated with the help of the evaluation unit.
In the next sections the processing steps of the system are described in more detail.

3.1 Search Request Specification Stage

The search request specification unit enables the user to generate a pictorial specification interactively from scratch or to invoke previously made specifications from the specification database for modification.
If no previously stored specification is suited, the user can select an example image from the example image database to help him or her to generate a pictorial specification. Regions of interest are interactively indicated by the user by putting a box around them, see fig. 2. The data in the box is copied to separate working pad. Then, the data is processed to transform the region of interest into a single detail with the aid of image processing functions and interactive graphical tools. The transformation is performed in order to obtain an optimal description of a search item in pictorial form which is used in the pictorial specification. The image processing and graphical procedures are working on the data in the working pad. They function in much the same way as a paint box, and are available through pushbutton commands.
The image processing procedures include: noise reduction (to obtain the essential shape of the data), edge detection and thresholding (to make the data two values: discriminating the detail from its surroundings), filter operations (to pre- and postprocess the data in the working pad), morphology operations and image component elimination facilities (to remove irrelevant image components from the detail), and some more.
The graphical tools include: wipe out facilities, drawing of lines, straight lines, polygones, circles, rectangles, splines, etc. (to introduce or modify pictorial elements in the detail).
The result of these operations is a detail of which the data represents a modified version of the original region of interest. The details are copied back at the place where the original region of interest has been indicated, see fig. 2.
If the user is ready generating details, the details are copied to the specification pad. The specification pad is used to place spatial relationships between the details (the straight white lines indicate relationships), yielding the pictorial specification, see fig. 3. Spatial relationships between the details include: one detail is north of an other detail, or north-east, east, etc, or one detail is always at a fixed distance of an other detail, variable distance, etc.

Figure 2: Example of the use of image processing routines and graphical tools.

Figure 3: Defining spatial relationships.

Hence, a pictorial specification is defined as a set of one or more details with their spatial relationships.

3.2 Matching Stage

The indexing unit performs the matching of the pictorial specification with every location in each of the images in the image database. During the acquisition process, noise and other distortions may be added to the images in the image database. Therefore, images in the image database are first preprocessed by noise reduction filters to obtain the essential shape of the data. After this process, the images are thresholded to obtain a binary representation of the grey-value image data. Pixels indicating the figure (foreground) receive a value +1 and the remaining pixels (background) receive the value -1. The matching operation is based on fuzzy Mathematical Morphology [2], [8], [9], in combination with a relationship controller. The use of fuzzy Mathematical Morphology permits some inconsistency and noise to be present in the data of the details as well as in the data of the images. To that end, the image data of the detail is transformed into a structering element in order to be able to apply the correspondence measure.

The transformation is as follows:

- Positive weights are given to foreground pixels directly proportional to the distance of the foreground pixels to the border.
  In this way, higher weights are given to foreground pixels further inwards so that foreground pixels at more certain locations receive higher weights than those which are positioned at less certain places (pixels located near contours, which may be due to noise)

- Pixels near the borders, both on the inside as well as on the outside, are given zero weights (they are the so-called 'don't care pixels'). This is done to permit a certain distortion in the model.

- Negative weights are given to background pixels.

The correspondence of a structering element and the local configuration of the data of the images is calculated by taking the point by point weighted sum of the structering element with the neighbourhood. On the locations where the pixels of the structering element and the corresponding pixels of the image neighbourhood have the same sign, the result of the operation is said to be a perfect HIT. If all signs are opposed, the result is a perfect MISS. Applying this operation to all locations of an image will yield a value of correspondence for every location in the image in the interval between the value of a perfect MISS and a perfect HIT. By determining the highest peak in the correspondence values (peaks arises from the fact that the fuzziness of the operation acts as an adoptive noise filter), the location of the best fit is derived. The relationship controller verifies whether the relationships between the details, as demanded in the search specification, hold for the image under investigation. For each image, the best fit (quality and position of the fit) and the pictorial specification are stored in the logical database. After the search, the logical database is ranked in descending order of quality of resemblance. The retrieved images are presented one by one to the user for interactive evaluation.

3.3 Interactive Evaluation Stage

The evaluation unit enables the user to visually judge the quality of the retrieval result. In this case, the evaluation unit enables the user to judge the quality of correspondence between the pictorial specification and the images visually by superposition of the original pictorial specification on the target images, see fig. 4.

Figure 4: Interactive evaluation.

At this point, the user may judge the results satisfactory or two situations occur:

- The system gives too few results.
- The system yields too many images which satisfy the search request, such implies that the search request is too weakly specified. By the visual presentation of the result, the user may learn how to adjust the pictorial specification to let it be more specific and makes a rerun.

4 Experiments

Experiments have been carried out with the image retrieval system on an existing database of more than 200 MRI images taken from the chest containing a variety of planar cross-sections through a large variety of patients. The images have been recorded at the Yale University Medical School facilities (who are partners in this project). To test the usefulness and performance of the image retrieval system, the following experiment was conducted.
A search specification was made from one of the images in the example image database. The pictorial specification consisted of three details with two spatial relationships among them. With this specification, it was the aim to find in the database those images which contained the same anatomical view. Generating the specification took approximately one minute of interaction. Thereafter, the program ran 200*40 seconds on a standard SUN-SPARC station. The result was that 6 out of the first 7 highest ranked images rightfully contained the desired plane (and 1 was not), whereas no planes where missing.

5 Conclusion

An image retrieval system has been presented in this paper, called Σnigma, that searches for pictorial objects in a database of raw images.
The system's pictorial approach is adopted from the observation that the majority of the pictorial information in an image cannot be fully expressed by textual or numerical information due to essential limitations in the expressive power of text and numbers. Therefore, the image retrieval system, in this paper, carries out a fully automatically searching for specific geometrical details with their relationships on a database containing raw images.
The system is characterized by the following capabilities:

- interactive search request specification ability.
- library building capability for pictorial specifications.
- the possibility to work on formalizable domains (e.g. documents of electronic schemas and geographical maps) as well as on weakly formalizable domains (e.g. medical images).
- robust to search requests containing slight distortions as well as noisy and distorted images.
- fully automatically retrieving (i.e. no abstracts are needed).
- interactive evaluation.

The image retrieval system has been implemented in C in combination with the X widget set, [1], on a SUN-SPARC workstation with UNIX as operating system. The ScilImage package, [8], provides the image processing functionality.

Acknowledgements

Prof. C. Jaffe and dr. M. Tagare are gratefully acknowledged from their help in this project.
The research project is supported by NIH: NLM-Ep(1 R01LM05007-01A1)

References

[1] Asente, P. J. and Swick, R. R.,
X Window Systems Toolkit,
Printice-Hall, Englewood Cliffs, New Jersey, 1989.

[2] Boomgaard van den, R.,
Threshold Logic and Mathematical Morphology,
Proc. 5th International Conference on Image Analysis and Processing, Positano, Italy, 1989, pp. 111-118.

[3] Chang, S.,
Image Information Systems,
Proc. of the IEEE, Vol. 73, No. 4, 1985, pp. 754-764.

[4] Grosky, W. and Mehrotra, R.,
Index-Based Object Recognition in Pictorial Data Management,
Computer Vision, Graphics and Image Processing, 52, 1990, pp. 416-436.

[5] Kate ten, T., Balen van, R., Smeulders, A.W.M., Groen, F.C.A., Boer den, G.,
SCILAIM: a Multi-level Interactive Image Processing Environment,
Pattern Recognition Letters 11, 1990, pp. 429-441.

[6] Lyengar, S. and Kashyar, R.,
Guest Editors' Introduction. Image Databases,
IEEE Trans. on Software Engineering, Vol 14, No. 5, 1988, pp.608-610.

[7] Maragos, P.,
Tutorial on Advances in Morphological Image Processing and Analysis,
Optical Engineering, No. 26(7), 1987, pp. 623-632.

[8] Serra, J.,
Image Analysis and Mathematical Morphology,
London, Academic Press, 1982.

[9] Smeulders, A.W.M.,
An Introduction to Image Processing and Computer Vision,
University of Amsterdam, 1991.

[10] Tamure, H. and Yokoya, N.,
A Management System for an Intergrated Database of Pictures and Alphanumerical Data,
Computer Graphics and Image Processing, No. 16, 1981, pp. 270-286.