# $\Sigma$nigma: An Image Retrieval System

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

The paper presents the functionality of a system which retrieves images on the basis of automatically generated indexes (i.e. semantic image representations, obtained by automatic image analysis, indicating the content of the images).
The system consists of two parts: an off-line indexing part and an on-line image retrieval part. The indexing component is used to automatically generate semantic representations of images so that the image retrieval component can use this information to enable image retrieval. The man-machine communication of the image retrieval component is based on an iconical graphical query language to accomplish geographical query specification for image access.
Experiments have been carried out on three different sets of images from the following domains: MRI images of the chest, electronic schemas and topographic maps. The experiments show encouraging results especially for domains which have a high degree of formality in their pictorial expression, such as electronic schemas and topographic maps, and to a less extent for domains having a weak degree of formality such as MRI images of the chest.

**keywords:** image indexing, image retrieval, search specification, graphical query language, fuzzy Mathematical Morphology.

## 1 Introduction

The growing interest of electronically storing digital images has led to a demand for information systems able to efficiently manage large amounts of images. The convergence of database, graphics and image processing/pattern recognition technology yields the basis for the creation of such image databases [4], [9]. Most of the existing image database systems are based on the paradigm to store images together with verbal descriptions to enable their retrieval [10], [12].
However, generating verbal descriptions for large amounts of images is almost always inadequate (i.e. the majority of pictorial information in an image cannot be fully captured by text and numbers due to the essential limitations in expressive power of language) and also time consuming and error prone. Therefore, an information system is proposed in this paper, called $\Sigma$nigma, capable of retrieving images on the basis of geometrical or iconical indexes.
The system also supports an iconical graphical query language to perform geographical query specification as opposed to current dbms-query languages which are far from expressing visual representations to depict the domain of interest [2], [5], [8].
In this paper, we first present the functionality of the system in Section 2. In Section 3 experiments carried out on three sets of images from different domains are discussed. Finally, a summary will be given.

## 2 Functionality of the System

$\Sigma$nigma is composed of an image indexing and an image retrieval component.
Image indexing is done off-line and consists of automatically generating semantic representations of images. From each image a set of feature instances is extracted by means of standard image analysis tools to represent the image, independent of a priori information about the content of the image.
Features may be defined domain independent or domain specific. Domain independent features may be based on generic geometrical properties such as peaks, pits and valleys or spatial derivatives as proposed by [7]. We will not discuss them at this point. Domain specific features are features from which the image is composed such as iconograms of components for electronic schemas or the legend signs for geographical maps. In this paper we focus on domain specific features.
Once the feature instances have been extracted, information about the degree of resemblance and position of the instances in image space is stored as a spatial index to enable retrieval.

### Indexing

The system enables the definition of domain specific features. Domain specific features are defined as salient pictorial patches and are generated by processing the data of example pictorial patches by image processing functions and graphical tools. Let $e(n, m)$ denote an example pictorial patch (which usually consists of 16x16 to 64x64 pixels), $b(n, m)$ its corresponding spatial binary image patch, $h$ an edge detector (currently of the Söbel type) and $i$ an automatic threshold function, then:

$$g = i \circ h$$

$$b(n,m) = g(e(n,m)) \quad (2.1)$$

The spatial binary image patch can be interactively processed by graphical tools to introduce or modify pictorial patch elements, see fig. 1.
Next, $b(n,m)$ will be processed to prepare it as a structering element for matching based on fuzzy Mathematical Morphology [3], [11]. Let $s(n,m)$ denote the corresponding structering element of $b(n,m)$, then:

$$s(n,m) = DT(b(n,m)) \quad (2.2)$$

Where $DT()$ is the Euclidean distance transform. Positive distance weights are given to structering element pixels directly proportional to their distance to the border. In this way higher weights are given to object pixels at more certain locations in the interior of the object. Pixels near the borders receive zero weigths (i.e. 'don't care pixels') and negative weights are given to background pixels.
The structering element $s(n,m)$ is used to find instances in unknown images similar to the patch it represents. A measure of correspondence between the structering element and all places in the images is expressed by taking the point by point weighted sum of the structering element and the local neighborhood of each image. Before performing the matching, each image $I(x,y)$ in the image database is preprocessed by the same image processing function $g$, which was applied on the example pictorial patch $e(n,m)$, to guarantee uniformity.
After this image preprocessing stage, image foreground pixels receive the value 1 and background pixels the value -1. The measure of correspondence is expressed by:

$$g = i \circ h$$

$$B(x,y) = \begin{cases} \textbf{-1}, & \text{if} \quad g(I(x,y)) = 0 \\ \textbf{1}, & \text{if} \quad g(I(x,y)) > 0 \end{cases}$$

$$r(x,y) = \begin{cases} \textbf{0}, & \text{if} \quad \sum_{n,m \in s} s(n,m)B(x,y) < t \\ \textbf{1}, & \text{if} \quad \sum_{n,m \in s} s(n,m)B(x,y) \geq t \end{cases} \quad (2.3)$$

A local correspondence of the binary image $B(x,y)$ with $s(n,m)$ is represented in the data field $r(x,y)$. Threshold $t$ determines spots in the image with high correspondence and is usually preset to $t = 80\%$ of the absolute sum over $s(n,m)$.

The system enables the domain expert to visually judge the quality of the matching operation by superposition of the original features on the target images at locations where corresponding feature instances have been found. The images are ordered in descending order of quality of resemblance, see fig. 1. In the case of too few correspondences, the domain expert can adjust threshold $t$ and make a rerun. When successful, the degree of resemblance and position of the instances are stored as indexes in the index database.

## Image Retrieval

$\Sigma$nigma allows for on-line image retrieval. The query language is based on semantic models where features and their relationships are conceptually represented by means of visual metaphors such as icons, diagrams and gestures to provide friendly man-machine query specification.
Although, it is generally impossible to anticipate all queries asked by the user, the system's flexibility enables the user to specify queries for which there is no index information available in the index database by combining icons representing already indexed features by means of (spatial) relationships yielding a search specification.
If the search specification consists of one item, retrieval is reduced by fetching the index information of that particular item for each image and to order the images in descending order of quality of resemblance. Once the images have been ordered, they are presented to the user with the superposition of the item on the target images in order to indicate where instances of that particular item has been found.
In the case that the search request is composed of more than one item, the user may specify relationships between them. Spatial relationships include: "N", "E", "S", "W", "NE", "SE", etc. and are represented by straight white lines, see fig. 2. Other relationships are "CONTAINED", "ATTACHED", "CLOSE". Feature attribute values are entered by simple dialogue boxes (e.g. the angle of a corner to be searched for).
After the user has specified relationships between the items, the search specification will be transformed into a graph (i.e. each node is a feature and edges denote relationships). Subgraph isomorphism matching is performed with the specification graph and the index information fetched for each image. In this way, instances, as specified by the search request, in the images are identified. After the graph matching operation, the images are ordered in descending order of quality of resemblance and presented to the user. The instances found are superpositioned on the target images in descending order of correspondence quality, see fig. 2.
By combining features and their relationships, object descriptions are obtained. In fact, a search specification consisting of more than one item and relationship is conceived as an object description. The user may store instances found by the system for objects, as specified by a search specification, as indexes in the index database. In this way, search specifications can

be generated which consist of semantically meaningful combinations of features and already specified and indexed objects. As a consequence, the data model supported by the system that represents the index information is object-oriented.

# 3 Experiments

Σnigma has been implemented in C in combination with the X widget set, [1], on a SUN-SPARC workstation with UNIX as operating system. The SciIImage package, [6], provides the image processing functionality.
A set of examples patches $e_j$, $j = 1...m$, together grasping domain specific information in spatial form and describing the basic spatial characteristics of the application domain, have been generated for two domains: electronic schemas and geographical maps. These two domains have a high degree of formality in their pictorial expression. The domain example patches $e_j$ correspond to signs in the legend for maps and to iconograms for electronic schemas and have been matched for each image in the image database. The degree of correspondence and position of the instances of $e_j$ have been stored as an index in the index database.
Tentative experiments have been carried out to test the usefulness and performance of the system. A search request have been specified to retrieve images containing four different electronic components with three spatial relationships among them from a small image database of 20 images. The number of domain specific features, $m$, from which the images were composed was 20. The system found all images in descending order of quality of resemblance containing the four components and relationships as specified by the search request. It took approximately 0.4 seconds per image on a standard SUN-SPARC station to fetch the appropriate index information from the index database and to perform graph matching with the specification graph and the index information.
A larger experiment has been carried out on an image 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. A search request was made which consisted of three domain specific features 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. Composing the request took approximately one minute of interaction. The result was that 6 out of the first 7 highest ranked images rightfully contained the desired plane (and 1 was not), whereas no instances were missed.

# 4 Summary

An image retrieval system has been presented in this paper capable of retrieving images on the basis of automatically generated semantic representations of images. The system is composed of an off-line indexing and an on-line retrieval component. In the indexing component, a set of feature instances is extracted from each image and stored as an index. Features are defined domain independent and domain specific. Finding instances of domain specific features in images is based on fuzzy Mathematical Morphology. Image retrieval is defined as a search request for images containing features as specified by the user by means of a graphical query language.
The system is characterized by the following capabilities:

- off-line interactive domain specification.
- automatically indexing of images (i.e. semantic image representations, obtained by automatic image analysis, assigned to each image indicating its content) to enable image retrieval.
- search request specification based on a graphical query language.
- automatically image retrieval.
- interactive evaluation.

Experiments show encouraging results for domains which have a high degree of formality in their pictorial expression (such as electronic schemas and topographic maps) and to a less extent for domains having a weak degree of formality (such as MRI images of the chest).


## Acknowledgements

Prof. C. Jaffe and dr. M. Tagare are gratefully acknowledged for 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] Batini, C., Catarci, T., Costabile, M. F. and Levialdi, S., **Visual Query Systems: A Taxonomy**, IFIP WG 2.6, 2nd Working Conference on Visual Database Systems, Budapest, Hungary, 1991, pp. 159-173.

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

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

[5] Haber, R. N. and Wilkinson, L., **Perceptual Component of Computer Display**, IEEE Comp. Graphics and Analysis, 1982, pp. 23-35.

[6] 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.

[7] Koenderink, J. J. and Doorn van, A. J., **Receptive Field Families**, Biol. Cybern., Vol. 63, 1990, pp. 291-298.

[8] Levialdi, S., **Cognition, Models and Metaphors**, IEEE Workshop on Visual Languages, Chicago, 1990, pp. 69-77.

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

[10] Salton, G. and McGill, M. J., **Introduction to Modern Information Retrieval**, McGraw-Hill, 1983.

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

[12] Tamura, H. and Yokota, N., **Image Database Systems: A Survey**, Pattern Recognition, Vol. 17, no. 1, 1984, pp 29-43.

Figure 1: Domain definition of electronic schemas a: Example image to extract salient image patches. b: image processing and graphical tools for processing the example pictorial patch. c: interactive evaluation of found instances.

Figure 2: Image retrieval of MRI images of the chest a: search request specification. b: interactive retrieval evaluation.