Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Knowledge-Based Systems 19 (2006) 687–695
Knowledge-Based SYSTEMS
www.elsevier.com/locate/knosys

# Iconic pictorial retrieval using multiple attributes and spatial relationships

Sam Y. Sung [1], Tianming Hu *

*Department of Computer Science, National University of Singapore, Kent Ridge, Singapore 119260, Singapore*

Received 19 January 2004; accepted 4 May 2006
Available online 4 August 2006

**Abstract**

This work is on the use of multiple attributes or features and spatial relationships, with the help of a user interface based on an iconic paradigm, to retrieve images represented by iconic pictures. An icon has texture, color, and text attributes. Texture is represented by three statistical textural properties, namely, coarseness, contrast, and directionality. For text, the vector space model is used. For color, a representation based on a modified color histogram method which is less storage-intensive is proposed. The final icon similarity is the combination of the attribute similarity values using a proven adaptive algorithm. 2-D strings and its variants are commonly used to represent spatial relationships and perform spatial reasoning. We extended the method to include similarity ranking by using different similarity functions for different spatial relationships and an efficient embedding algorithm. Furthermore, our method solves the problem of query expressiveness which all methods based on 2-D string representations suffer from.


*Keywords:* Content-based indexing; Information retrieval; Pattern recognition; Knowledge base system; Image database

## 1. Introduction

An iconic picture can be viewed as a logical picture, while the image it represented can be viewed as a physical picture. The logical picture representations are high-level abstracted representations, denoting the logical relationships among picture objects or icons [5]. Each icon in our work is bounded by its MBR (minimum bounded rectangle) and contains additional attributes such as texture, color and text which can be automatically computed or manually specified. The use of multiple attributes is important for effective retrievals because each different single attribute retrieves a different subset of relevant images. Furthermore, the combination of content-based attributes (texture and color) and text offers a richer way of describing an image object represented by an icon, for text alone is unable to describe an image fully.

The use of logical pictures obviates the need for repeated image understanding. Functions for spatial similarity retrieval based on symbolic images are useful in distributed environments where physical images are stored at separate image stores while symbolic images are stored at each local site. Only those images that are relevant by comparison of their logical pictures, are transferred from the image stores to the local site [10]. Also, with this separation, many logical pictures may represent a single image. This is important because the same image may be used in different ways during different time periods.

A visual query system can be categorized into four main paradigms, namely tabular, diagrammatic, iconic and hybrid [2,1,4,17]. Our work is based on the iconic paradigm and our contribution is as follows:

* Corresponding author. Address: Department of Computer Science, DongGuan University of Technology, DongGuan, GuangDong 523808, China.

*E-mail addresses:* sysung@southtexascollege.edu (S.Y. Sung), tmhu05@gmail.com (T. Hu).

[1] Department of Computer Science, South Texas College, McAllen, TX 78501, USA.

0950-7051/$ - see front matter 
doi:10.1016/j.knosys.2006.05.013

(1) Visual query systems use the full power of new technologies, like two/three dimensional representations, colors, multiple windows, etc., thus extending the man-machine communication bandwidth in several directions [2].
(2) Our implementation is web-based, where most users are non-professional. The iconic paradigm is intuitive and next-to-effortless to learn and caters for this group of users.
(3) Useful when the user is not informed about the semantic content of the database in advance (very likely in our web-based implementation) since the visibility of the icons paints a rough semantic content of the database.
(4) Can be used to provide a uniform interface for both image registration and querying subsystems. This paradigm has some shortcomings such as possibility of icon ambiguity and icon overcrowding. Icon classes are arranged in an inheritance hierarchy to alleviate the icon overcrowding problem.

Our system is divided into two main subsystems: image registration (IRS) and query (QS). In IRS, users instantiate icons from appropriate icon classes by marking their MBRs over possible image objects on the current image to be registered with the database. In QS, users instantiate icons from icon classes and specify spatial relationship constraints between any two instantiated icons to pose a query.

## 2. Related works

Text retrieval techniques can be divided into exact match and partial match [14]. Exact match misses many relevant documents that match partially and also does not rank the result. We use the vector space model, a partial match technique where both the query and documents are represented by vectors of term weights $(d_1, d_2, \ldots, d_m)$. A commonly used similarity function between the query and document vectors is the cosine correlation [15]. Texture analysis has been widely studied and many approaches have been developed. Some of the common approaches are statistical, such as Wold transform and Gabor filters. We used the statistical texture properties coarseness, contrast, and directionality which are obtained by computing the local statistical distribution of image intensity and are defined in [22]. The similarity functions used to compare each of the properties are different from those suggested in [22] and will be shown later. Color distribution and color spatial are two main approaches for representing colors [23]. A color distribution is defined by a color histogram [12]. Two histograms $H_q$ and $H_t$ can be compared using the Manhattan distance or other suitable distance measures. Histograms are fast but do not store spatial information; two images having the same color distribution may look different. Further, they are sensitive to background noises and lighting conditions [20]. Our method is based on the color distribution approach but is less storage intensive compared to normal histograms. Color spatial techniques are known to perform better than color histograms; some examples of this approach are mentioned in [20,11].

Spatial relationships are also intensively researched in the image retrieval literature. Gudivada and Raghavan [10] partitions spatial queries into two categories: spatial similarity approach (ranks results based on a similarity ranking function), spatial exact match approach (requires an algorithm that provides a Boolean response). A common exact match technique is the use of 2-D strings and its variants [7,8,6,18]. For this method, different types of queries with increasing restrictions are defined: Type-0 (existing), Type-1 (category), Type-2 (orthogonal), and Type-3 (coordinate) [18]. Extensive examples of the spatial similarity approach can be found in [10,19]. Neither method in [10] or [19] is able to do spatial reasoning, such as finding an object which "overlaps" another object, while Type-*i* approach cannot produce a ranked list of results since no effective similarity value is computed. For the best of both worlds, we integrated the two approaches.

Integrating similarity computation into the Type-*i* method using the 2-D string representations may not be effective since this representation suffers from the problem of inadequate resolution, i.e., only a few discrete similarity levels are returned [10]. In order to solve this problem, we represent each icon's position in the picture by its MBR. Spatial relationships are implicit in this representation and spatial reasoning can be performed. In [10], only a single spatial similarity function based on orientation is used which is insufficient because for certain relationships, same orientations may not mean similar spatial relationships according to human perception. Also it is not very effective for non-point objects. Fig. 1 illustrates this. We, therefore, defined four different spatial similarity functions for the disjoint, meet, contain, and overlap relationships in our work.

All methods based on 2-D string representations suffer from the problem of query expressiveness which arises because the user cannot specify the relationship type (existing, category, etc.) at the icon level (between any two icons) but only at the query level (applied to the whole query). This scheme is either too relaxed or too restrictive for certain queries as shown in Fig. 2. Our work solves the query expressiveness problem by allowing the user to specify relationship types between any icons in the query.

Finding a sub-picture in an iconic picture that matches the query to some degree of similarity is almost akin to

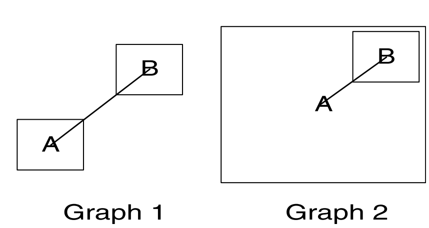


Fig. 1. Since graphs 1 and 2 have the same centroid orientations, the method in [10,2] will treat them to be exactly similar. However, they do not look similar.