R-Histogram: Quantitative Representation of Spatial Relations for Similarity-Based Image Retrieval*

Yuhang Wang
Department of Computer Science
Dartmouth College
Hanover, NH 03755, USA
wyh@cs.dartmouth.edu

Fillia Makedon
Department of Computer Science
Dartmouth College
Hanover, NH 03755, USA
makedon@cs.dartmouth.edu

ABSTRACT
Representation of relative spatial relations between objects is required in many multimedia database applications. Quantitative representation of spatial relations taking into account shape, size, orientation and distance is often required. This cannot be accomplished by assimilating an object to elementary entities such as the centroid or the minimum bounding rectangle. Thus many authors have proposed numerous representations based on the notion of histograms of angles. However, they can only represent directional relations, but not the topological spatial relations "inside" and "overlap." Moreover, distance information is not explicitly taken into account. To address these issues, we propose in this paper a new histogram representation called R-Histogram that extends the histogram of angles by incorporating both angles and labeled distances. Dissimilarity between images is then defined by the distance between corresponding R-Histograms. A prototype Query By Example (QBE) system using the R-Histogram has been implemented. The effectiveness of our algorithm is demonstrated with experiments on two databases of 2000 synthetic images.

Categories and Subject Descriptors
H.3.1 [**Information Storage And Retrieval**]: Content Analysis and Indexing; H.2.8 [**Database Management**]: Database Applications—*Image databases*

General Terms
Design, Management

Keywords
Spatial Relations, R-Histogram, Similarity Search, Image Retrieval

*This work was supported in part by National Science Foundation under grant no. IDM-0083423.

1. INTRODUCTION
There are many general-purpose content-based image retrieval systems, e.g. the QBIC [6] system and the Photobook [14]. They mainly use color, texture and shape as image features. However, representing the spatial relations between objects is also an important component of image content description and access. For example, the spatial relationship between brain lesions and anatomical brain structures in medical images is critically important for early disease diagnosis and thus important for image retrieval. Typical applications of spatial relation representations are content-based image retrieval (e.g. [3, 8, 15, 17]), video indexing and retrieval (e.g. [5]), computer vision, robot navigation, and Geographic Information Systems (GIS). To assess the degree of similarity of two images according to the spatial relations between objects, first we need to extract a compact representation of spatial relations from images, and then define a (dis)similarity measure (e.g. a distance function) on such representations. Our ultimate goal is to answer queries like "find similar MR images to one with a lesion inside the frontal lobes", or "find similar surveillance video sequences to one in which a man walks from the middle of a room to the east side."

In this paper, we propose a new histogram representation of spatial relations called R-Histogram. We assume the images are segmented and each object is assigned a unique label; i.e., we deal with symbolic images, as defined formally in [8]. The dissimilarity between two images is then defined by the distance between the two corresponding R-Histograms. We have implemented a prototype Query By Example (QBE) system using the R-Histogram and tested the effectiveness of our algorithm on two databases of 2000 synthetic images.

The remainder of the paper is organized as follows. We discuss related work in the next section. In Section 3, the notion of R-Histogram is introduced. In Section 4, we describe the experiments we have performed and provide results. We conclude in Section 5.

2. BACKGROUND AND RELATED WORK
Significant work has been reported on spatial relation representation. Many authors have stressed the importance of qualitative spatial relationships [4]. Approaches have been based on Allen's interval relations [1] (e.g. [13, 16]), 2D strings [3] and their variants, Attributed Relational Graphs (ARGs) (e.g. [15]) or the spatial orientation graph (e.g.

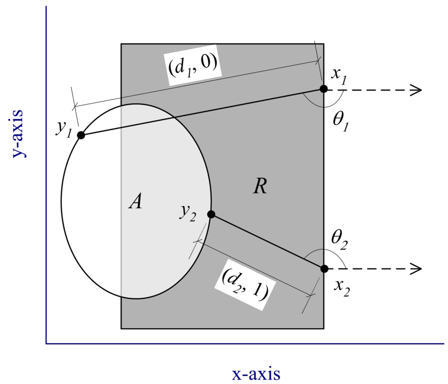


Figure 1: Examples of the angle $\theta(x, y)$ and the labeled distance $LD(x, y)$. The gray rectangle object R is the reference object and the white ellipse object A is the object of interest. d_i is the Euclidean distance between x_i and y_i.

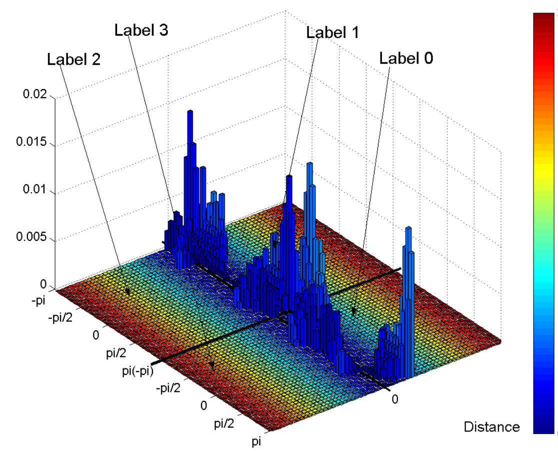


Figure 2: $RH(A, R)$ for the two objects in Figure 1. Each quadrant is associated with a unique label.

Table 1: Labels in the Labeled Distance

Pixel x inside A	Pixel y inside R	$l(x, y)$
False	False	0
False	True	1
True	False	2
True	True	3

[8]). All of these approaches assimilate an object to very elementary entities such as the centroid (e.g. [8, 9]) or the minimum bounding rectangle (e.g. [13]). This simplification process cannot give a satisfactory modelling of the spatial relations. For example, projecting two objects to each of the dimensions and considering each dimension separately is inadequate, because the two objects may not overlap at all when their projections onto the x and y axes overlap simultaneously.

Freeman [7] proposed that the relative position of two objects be described in terms of 13 primitive spatial relations: 1) LEFT OF, 2) RIGHT OF, 3) ABOVE, 4) BELOW, 5) BEHIND, 6) IN FRONT OF, 7) BESIDE, 8) NEAR, 9) FAR, 10) TOUCHING, 11) BETWEEN, 12) INSIDE, and 13) OUTSIDE. The first six are called the primitive directional relations (See [2] for a review and comparison of directional relative position representations). In [12], Miyajima and Ralescu introduced the notion of the histogram of angles to represent directional relations. Given an object of interest A and a reference object R, the histogram of angles is computed from the angle between any two points in both objects and normalized by the maximum frequency. This histogram represents the spatial directional relations of the object A with respect to the reference object R. Numerous studies are based on this notion of histogram of angles, e.g. [9, 10, 12]. In [11], Matsakis et. al proposed the concept of the Force Histogram, which considers pairs of longitudinal sections instead of pairs of points. However, all these angle histogram approaches can only represent directional relations, but not the topological spatial relations "inside" and "overlap". For instance, it is not reasonable to conclude that an object A is inside another object B, knowing that B is somewhat above, below, to the right, and to the left of A as well. Moreover, distance information is not explicitly taken into account and it is a real disadvantage. In many cases, quantitative representation of spatial relations taking into account shape, size, orientation and distance is required.

In summary, angle histogram approaches provide quantitative directional relation representation, but are not ideal for domains such as medical content-based image retrieval systems because:

1. They don't provide explicit metric (distance) information that is important in these domains;

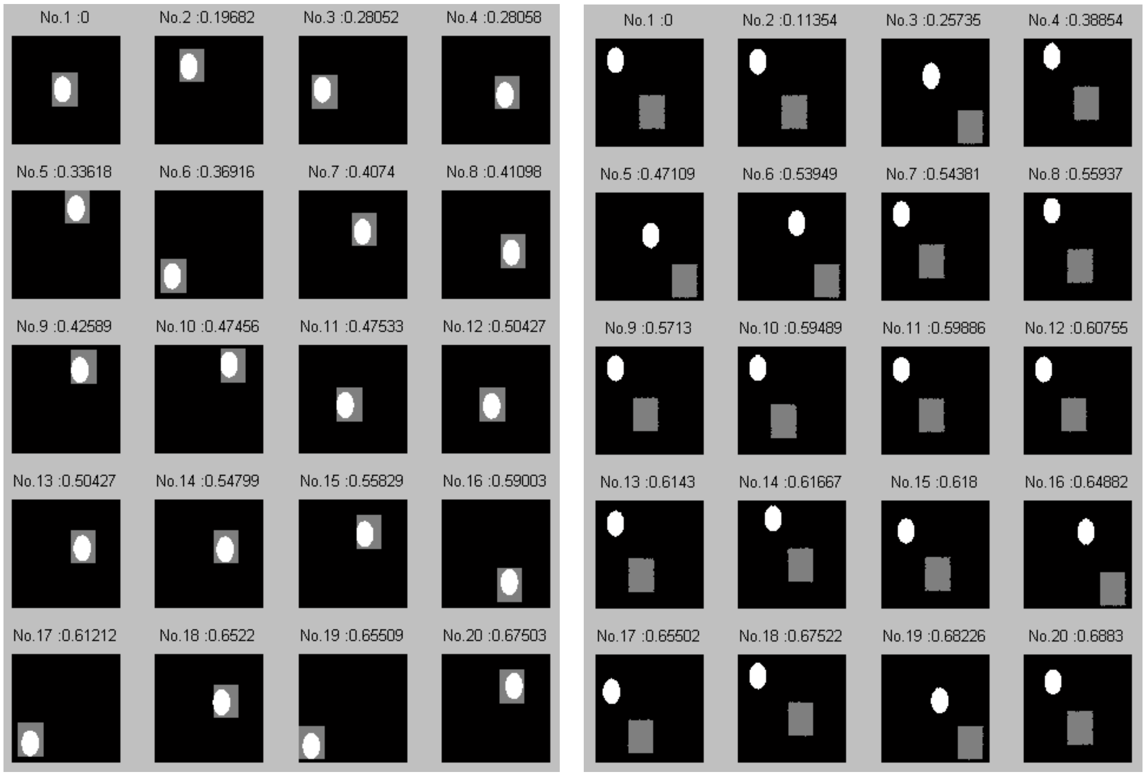
2. They don't support extraction of topological spatial relations like "inside", "overlap".

The R-Histogram extends the angle histogram by incorporating labeled distances to address these issues, as described in the next section.

3. HISTOGRAM REPRESENTATION

3.1 The R-Histogram

Given a reference object R and an object of interest A, the goal is to represent, quantitatively, the spatial relations between R and A. Consider the vector originating from a pixel x on the boundary of R to a pixel y on the boundary of A. If x and y don't coincide, we compute the angle between the x-axis of the coordinate frame and $\overrightarrow{xy}$. This angle, denoted by $\theta(x, y)$, takes values in $[-\pi, \pi]$. As in histogram of angles [12], the set of angles from any pixel on the boundary of R to a pixel on the boundary of A expresses the directional relations between R and A. The novel idea introduced in this paper is the labeled distance. The labeled distance from x to y, denoted by $LD(x, y)$, is defined as a pair $(d(x, y), l(x, y))$, where $d(x, y)$ is the Euclidean distance from x to y and $l(x, y)$ is defined in Table 1. Here, column 1 describes whether pixel x is inside object A, and column 2 describes whether pixel y is inside object R.



(a) (b)

Figure 3: The results of querying the databases with example images. (a) and (b): results obtained from the image database A and B respectively. In each case, the No.1 image is the query image and the best 20 matches are shown. The images are ordered from left-to-right and top-to-bottom in increasing distance from the query image, with the distance following the rank shown above each image.

For the set of vectors originating from any pixel on the boundary of R to any pixel on the boundary of A, we construct a histogram as follows: Let x and y be the pixels on the boundary of R and A respectively. The bin $H(I,J,L)$ is incremented as follows:

$$H(I,J,L) = \begin{cases} H(I,J,L)+1 & \text{if } \theta(x,y) \in A_I \\ & \quad \wedge\ (d(x,y) \in D_J \\ & \quad \wedge\ l(x,y) = L \\ H(I,J,L) & \text{otherwise} \end{cases} \qquad (1)$$

where A_I is the range of angle values spanned by bin $H(I,J,L)$, D_J is the range of distance values spanned by bin $H(I,J,L)$, and $L \in \{0,1,2,3\}$ is the label associated with the distance values spanned by bin $H(I,J,L)$.

Then the histogram is normalized as follows:

$$h(I,J,L) = \frac{H(I,J,L)}{\sum_{I'=1}^{n_A} \sum_{J'=1}^{n_D} \sum_{L'=0}^{3} H(I',J',L')} \qquad (2)$$

where n_A is the number of angle bins and n_D the number of distance bins. The normalized histogram, denoted as $RH(A,R)$, is defined to be the R-Histogram of object A relative to object R.

A R-Histogram example is illustrated in Figure 2, where the x-axis is associated with angles and the y-axis with distances. Each of the four quadrants shows a portion of the histogram with one of the four labels associated with the distances. Clearly, $RH(A,R)$ is invariant to translation: when a translation is applied to A and R, the histogram does not change. When a rotation is applied to A and R, $RH(A,R)$ is simply shifted along the x-axis. $RH(A,R)$ can also be scale-invariant if the distances are normalized by dividing them by the maximum distance. $RH(A,R)$ encapsulates rich information about A and R. It is sensitive to their shape, orientation, distance and topological spatial relations. For example, if A is inside R, then the histogram will only be populated in quadrant 2.

Let N be the number of pixels in an image. We assume the objects are homeomorphic to a 2-ball. In the worst case, the number of pixels on the boundary of an object is $O(N)$. Therefore, the computation of R-Histogram takes $O(N^2)$ time. If the objects are convex, the number of boundary pixels will be $O(\sqrt{N})$ and the time complexity will drop to $O(N)$.

3.2 Distance Metric

The dissimilarity between two images is defined by the distance between corresponding R-Histograms. There are many histogram distance metrics. The distance metric used here is the histogram intersection. It is shown in [18] that when the histograms are normalized, the histogram intersection is

Table 2: Retrieval Performance

Methods	Precision
R-Histogram with labeled distance	0.92
R-Histogram with unlabeled distance	0.83

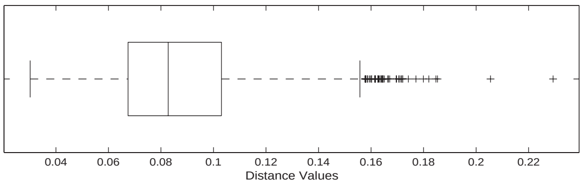


Figure 4: Segmentation error robustness evaluation. The box has lines at the lower quartile, median, and upper quartile.

given by

$$D(h_1, h_2) = \sum_{I=1}^{n_A} \sum_{J=1}^{n_D} \sum_{L=0}^{3} |h_1(I, J, L) - h_2(I, J, L)| \quad (3)$$

4. EXPERIMENTAL RESULTS

We have implemented a prototype QBE system using the R-Histogram in MATLAB and the C programming language. In this section, we present some examples to illustrate the effectiveness of the R-Histogram for image retrieval. The database A used for our experiments consists of 2000 synthetic images of size 256x256 containing rectangle and ellipse objects. The locations of the objects are random, rendering all possible kinds of relative spatial relations. The database B is created by adding random Gaussian noise to objects in each image in database A. In the experiments, we used R-Histograms with 36 angle bins and 50 distance bins. Generating the R-Histogram for each image takes about 19 seconds on a Mobile Pentium III 850MHz PC with 512MB RAM. This step is performed offline only once.

For each database, we tested 20 sample images chosen randomly from the database. Only two examples are shown in Figure 3 due to limited space. Based on human's relevance judgment, we obtained the precision values as shown in Table 2. Recall was not calculated because the database is large. We compare the results with those obtained with unlabeled distance. It is clear that R-Histogram with labeled distance outperforms that with unlabeled distance.

In order to study the robustness of our proposed method against segmentation error, we compute the distance between the R-Histograms for each image in database A and the corresponding noisy image in database B. The box plot for the 2000 distance values is shown in Figure 4.

5. CONCLUSIONS AND FUTURE WORK

The main contribution of this paper is to describe a simple and compact histogram representation that encapsulates directional and topological spatial relations between two objects. Even with such simple bin-to-bin histogram distance metric as histogram intersection, the proposed R-Histogram gives satisfactory similarity search results. To model the spatial relations of *multiple objects* in an image, we can use R-Histograms as the arc attributes in ARGs. The work presented in this paper is the first step in developing a similarity search system for a large time series image database. Our next goal is to improve the time complexity of R-Histogram computation and investigate the possibility of extracting semantic meanings from the R-Histogram representations. This would facilitate searching with descriptive key words. Future work also includes: evaluate the performance of more sophisticated cross-bin histogram distance metrics, extend the 2D implementation to 3D, add relevance feedback mechanism, and investigate time series query.

6. REFERENCES

[1] J. F. Allen. Maintaining knowledge about temporal intervals. *Commun. ACM*, 26(11):832–843, 1983.

[2] I. Bloch and A. Ralescu. Directional relative position between objects in image processing: a comparison between fuzzy approaches. *Pattern Recognition*, 36(7):1563–1582, 2003.

[3] S.-K. Chang, Q.-Y. Shi, and C.-W. Yan. Iconic indexing by 2-D strings. *PAMI*, 9(3):413–428, 1987.

[4] A. G. Cohn and S. M. Hazarika. Qualitative spatial representation and reasoning: an overview. *Fundamenta Informaticae*, 46(1-2):1–29, 2001.

[5] S. Dagtas and A. Ghafoor. Indexing and retrieval of video based on spatial relation sequences. In *Proc. of ACM Multimedia'99*, pages 119–122, 1999.

[6] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by image and video content: the QBIC system. *Computer*, 28(9):23–32, 1995.

[7] J. Freeman. The modelling of spatial relations. *Computer Graphics and Image Processing*, 4(2):156–171, 1975.

[8] V. N. Gudivada and V. V. Raghavan. Design and evaluation of algorithms for image retrieval by spatial similarity. *ACM Trans. on Information Systems*, 13(2):115–144, 1995.

[9] J. Keller and X. Wang. Comparison of spatial relation definitions in computer vision. In *Proc. of ISUMA - NAFIPS'95*, pages 679–684, 1995.

[10] R. Krishnapuram, J. Keller, and Y. Ma. Quantitative analysis of properties and spatial relations of fuzzy image regions. *IEEE Trans. on Fuzzy Systems*, 1(3):222–233, 1993.

[11] P. Matsakis and L. Wendling. A new way to represent the relative position between areal objects. *PAMI*, 21(7):634–643, 1999.

[12] K. Miyajima and A. Ralescu. Spatial organization in 2D segmented images: Representation and recognition of primitive spatial relations. *Fuzzy Sets and Systems*, 65(2-3):225–236, 1994.

[13] M. Nabil, A. H. H. Ngu, and J. Shepherd. Picture similarity retrieval using the 2D projection interval representation. *IEEE Trans. Knowl. Data Eng.*, 8(4):533–539, 1996.

[14] A. Pentland, R. W. Picard, and A. Sclaroff. Photobook: Content based manipulation of image databases. *Int'l J. of Computer Vision*, 18(3):233–254, 1996.

[15] E. Petrakis, C. Faloutsos, and K.-I. Lin. ImageMap: an image indexing method based on spatial similarity. *IEEE Trans. on Knowl. and Data Eng.*, 14(5):979–987, 2002.

[16] J. Sharma and D. M. Flewelling. Inferences from combined knowledge about topology and directions. In *Advances in Spatial Databases*, volume 951 of *Lecture Notes in Computer Science*, pages 279–291. 1995.

[17] C.-R. Shyu and P. Matsakis. Spatial lesion indexing for medical image databases using force histograms. In *Proc. of IEEE CVPR'01*, pages 603–608, 2001.

[18] M. Swain and D. Ballard. Color indexing. *Int'l J. of Computer Vision*, 7(1):11–32, 1991.