Orientation Radiograms for Image Retrieval: An Alternative to Segmentation

J. Bigün S. K. Bhattacharjee S. Michel
Signal Processing Laboratory
Swiss Federal Institute of Technology
CH-1015 Lausanne, Switzerland
Email: {*josef, sushil, michel*}*@ltssg4.epfl.ch*

(C) In Proc. 13th Int. Conf. Pattern Recognition*, Vienna, Austria, Aug. 25 – 30, 1996*

Abstract

For content based image retrieval using shape descriptors, most approaches so far extract shape information from a segmentation of the image. Shape features derived based on a specific segmentation are not suitable for images containing complex structures. Further, static segmentation based approaches are useful only for a small set of queries. In this paper we discuss the limitations of such boundary based shape features, and propose an alternative shape characterization technique based on orientation radiograms. A working image retrieval system based on this approach is described, and sample results are presented for a *full-image* query.

1 Introduction

The subject of content-based image retrieval has received enormous attention from the image processing community in recent years. Image retrieval systems typically administer large collections of images, and allow the user to retrieve a set of images which are relevant to a given query. Initial efforts in this direction consisted of extending standard database systems, usually relational systems, to handle images as well. The file name of the image was one field in a record. Other fields in the record contained some attributes of the image. These attributes were extracted manually. The main limitation of such a system is that description of an image is fixed, and reflects only the perception of the person who created the description. Retrieval systems based on such a rigid set of attributes can service only a small set of queries.

Content-based image retrieval (CBIR) systems address the issue of automatic indexing and retrieval of images. Descriptive features are automatically extracted from an image. The relevance of an image in a collection, to a given query, is automatically computed based on the degree of match between the set of features present in the image, and that present in the query. Since automatic feature extraction can be a computationally expensive process, the usual approach is to save pre-computed features and to use them during retrieval.

Several approaches have been proposed for CBIR. The most commonly used features, for collections of still images, are based on color and shape information. In the approaches that have been proposed so far, image retrieval based on shape descriptors requires a segmentation of the image. Shape descriptors are then computed for the segmented regions. Besides the well understood difficulties in generating ideal segmentations, this approach also has the following problem. The segmentation of an image is static, and is fixed at the time of *data population*, that is, when the image is added to the collection.

In this paper we put forward the contention that segmentation based approaches are not best suited for image retrieval problems. We suggest an alternative approach, based on *orientation radiograms*, where images are characterized using low-level features that can be easily used for comparisons with queries.

In section 2 we discuss the limitations of segmentation based approaches in more detail. The basic notions of orientation radiograms are introduced in Section 3. Orientation radiograms have been successfully used for image retrieval. The theoretical aspects of such a system are discussed in Section 4 and an operational system based on these concepts is described in Section 5. In section 6, we outline ideas for extending the existing system to handle *sub-image* queries as well, based on length-filtered orientation radiograms. Some conclusions are presented in Section 7.

2 The Problem with Segmentation

Several operational CBIR systems offer the facility of shape-based retrieval of images[1, 2]. The approaches proposed so far for characterizing shape, involve an automatic segmentation of the image. Assuming that a proper segmentation is available, the next step is to generate a shape description for each region. The shapes are usually described in terms of the boundaries of the regions. For example, QBIC uses a set of 20 moments to characterize shape [1]. Other boundary based shape features are derived from the *interesting points*, the points of maximum curvature, of the boundary [2].

Figure 1 shows a gray level image containing a complex structure. This image is rich in shape information that can be exploited for the purposes of image retrieval. It is intuitively obvious, however, that boundary based shape descriptors will not be adequate for characterizing such images for the purposes of CBIR. One might be inclined to argue that a large enough set of moments can be used to the shape information of complex shapes. However, it can be easily shown that higher order moments emphasize the outer boundaries more than the internal complexities. Another, and more significant problem is that moments are global features and are not additive in nature. Thus, they are not useful for retrieval based on sub-regions of images, where the user is interested in images containing the query-image as a part of a larger image.

Figure 1: An *ornament* image. Shape based retrieval will fail for such images if the indexing features are derived from a segmentation of the image.

The most severe problem with segmentation based approaches is that the decision as to whether a feature in an image is important is made at the time of data population. The segmentation is static, and subsequent queries do not have any influence on it. Picard [3] suggests pre-computing multiple segmentations for an image, each reflecting a different aspect of the image. However, even when it is possible to segment an image in different meaningful ways, there is no way of automatically generating a set of segmentations that encompass all possible interpretations of an image. *Thus, while automatic segmentation does offer a solution to the problem of manual image description, the underlying problem in the initial relational database oriented methods – that the high level of abstraction is too drastic a decimation of information, and can be useful to only a small set of queries – still remains unsolved.*

In this paper, we describe an approach to image indexing and retrieval that is based on shape information but does not rely on segmentation. An image is described in terms of the linear structures present in it. The description is based on a set of orientation radiograms, which contain information about the linear structures oriented in a specific direction.

3 Orientation Radiograms

The process of deriving the orientation radiograms of an image is presented in this section. The image is first decomposed into its *iso-orientation components*, also called *orientation images*. Each orientation image of the decomposition can be seen as the response to a (not neccesarily linear) filter tuned to a specific orientation, called the *pass orientation*. Thus, an orientation image contains information about the linear structures in the original image which are oriented along the pass orientation of the filter. Summing the responses of the orientation image along pass orientation yields a one dimensional projection, which is called an *orientation radiogram* [4]. (The name is derived from an analogy with X-ray imaging.)

To illustrate the representation of shape using orientation radiograms, we consider the decomposition of the synthetic image of a rectangle into six orientation images. The orientation images considered in our experiments have one of the following six pass orientations: 0^o, 30^o, 60^o, 90^o, 120^o, and 150^o. Figure 2 represents the (hypothetical) results of this decomposition. Only two orientation images, corresponding to pass orientations of 0^o and 90^o, have non-zero responses. The orientation radiograms corresponding to the decomposition of Fig. 2 are shown in Fig. 3. The one-dimensional orientation radiograms can be used to characterize a given shape, and for shape based discrimination.

Orientation radiograms can be obtained efficiently using linear symmetry computation [4]. For each pixel of the input image, a *linear symmetry vector* is computed. The linear symmetry vector, z, is given by

$$z = \left(\nabla\vec{\mathbf{f}}\right)^2 * m \ , \qquad (1)$$

where $\nabla\mathbf{f}$ and $*m$ represent the complex gradient image, $\mathbf{f}_x + i\mathbf{f}_y$, corresponding to an image $\mathbf{f}$, and convolution with an

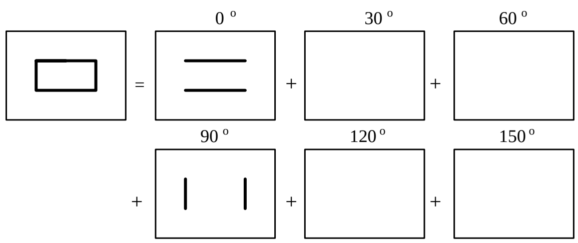

Figure 2: Decomposition of the image of a rectangle into six preselected orientations. The orientation corresponding to the decomposition is marked above the box.

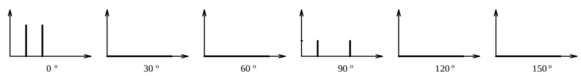

Figure 3: Orientation radiograms corresponding to the decomposition shown in Fig. 2. Each radiogram is obtained by projecting the corresponding orientation image along its pass orientation. The pass orientation of each radiogram is marked below the plot.

averaging filter, m, respectively. In geometrical terms, the linear symmetry vector corresponds to optimal straight-line fitted to the local power spectrum. The orientation of this vector gives the dominant orientation of the local neighborhood. The magnitude is interpreted as a measure of certainty of this estimation.

Using the linear symmetry vector, we can compute the orientation images, $r_l(\vec{x})$, as follows:

$$r_l = |z| \cos^2(\theta_l - \varphi), \tag{2}$$

where φ gives the orientation of the linear symmetry vector, z at a pixel-position, and θ_l specifies the pass orientation. In our experiments, $\theta_l = \frac{l\pi}{6}, l = 0, 1, \cdots, 5$. Therefore, we obtain six orientation images for each image. The length $|z|$ is normalized linearly, so that it lies in the real range $[0, 1]$. In r_l, the maximum pixel-value, $|z|$, occurs when $\theta_l = \varphi + n\pi$, and minimum, 0, occurs when $\theta_l = \varphi + \frac{\pi}{2} + n\pi$. As the difference between φ and θ_l increases, the response (i.e., pixel-values in r_l) decreases. This decrease cannot be regulated in the framework of Eqn. 2. We use the following definition for orientation images:

$$r_l(\vec{x}) = |z| \left(\exp(\beta \cos^2(\theta_l - \varphi)) - 1\right) \ , \tag{3}$$

where the parameter β can be used to control the orientation selection sensitivity. In our experiments, a constant value of $\beta = 2.4$, has been used for all the images in the collection.

4 Image Retrieval Using Orientation Radiograms

In this section we briefly describe the procedure of retrieving images from a collection, based on orientation radiograms. Before discussing the retrieval procedure, the criterion for comparing images is described.

Radiograms are compared based on their Fourier coefficients. An ordered set of ten Fourier coefficients, excluding the DC coefficient, is used to represent a radiogram. The peaks (maxima) of the radiograms are the most salient features. Using the first ten harmonics seems to preserve the peaks quite well. The Euclidean distance between two sets of Fourier coefficients is used as a measure of similarity between the corresponding radiograms.

When a new image is added to the collection, six orientation radiograms are computed for the image. A vector composed of the first ten Fourier harmonics is constructed for each radiogram. Thus we have six vectors of Fourier coefficients for each image in the collection. (In order to make the system invariant to image *flip*, we consider only the real part and the absolute value of the imaginary part, of each (complex) Fourier coefficient.)

A query is specified in the form of an image. For retrieval, the six radiograms are computed for the query image. As with the images in the collection, the radiograms of the query image are represented by their respective 10-D vectors of Fourier coefficients. The set of radiograms of query image is compared with that of each image of the collection, in turn. The similarity between each radiogram of the image and the radiogram of corresponding orientation of the query image, is computed. This results in a set of six similarity measures for a given image of the collection. The sum of these six similarity values (Euclidean distances) is considered to be a measure of similarity between the query and the image. The higher the number, the less similar the image is, to the query. The images in the collection may be ranked according to their similarity to a given query.

5 An Image Retrieval System for Ornament Images

An image retrieval system using orientation radiograms described above has been successfully implemented for a collection of *ornament images*, an example of which is shown in Fig. 1. The collection contains 500 gray level images of 94 ornaments. Each image in the collection contains one ornament. Images of the ornaments were obtained from old books in the library at the University of Lausanne [5]. One hundred images of ornaments were scanned. Since the books were old and odd sized, the ornaments were first photo-copied and then scanned as gray level images.

This process of course introduces further noise in the images. It was found that in the case of six pairs of images, both represented the same ornament. Thus the entire set of images represents only 94 distinct ornaments. This initial set has been expanded into a collection of 500 images by applying five different combinations of transformations such as horizontal/vertical flip, small amounts of rotation, and different amounts of gamma correction, to each of the 100 original images.

We present the retrieval results for one query. The query is the image of an ornament which has been rotated by 5^o, and flipped horizontally. Further a gamma-correction with $\gamma = 0.4$ has been performed on this rotated, flipped image. Figure 4 shows the query, and the five images that the system adjudged to be the most similar to the query. The image on the top-left corner shows the query. The image on the top-right is the 'closest' image in the collection to the query. Note that the first four images retrieved by the system all represent the same ornament. Only the last image (bottom-right) represents a different ornament. This is only because the collection contains only four images of the ornament specified in the query. Since a ranked list of the five images was requested, the system is forced to retrieve an image of a different ornament. In all our experiments, there was no false rejection when a list of five most similar images was requested. If only the closest match to the query is considered, the system shows an accuracy of 97.8% when radiogram based shape descriptors are used. By comparison, the accuracy of the same system falls to 49% when Reddi's moment-based shape descriptors are used as indexing features [6]. Of course, these performance numbers are relevant only from the point of view of a pattern classification problem.

Figure 4: Results for a specific query. **Top left**: Query image. The other images are the ranked list of images returned by our retrieval system (ranked by similarity to the query), for this query. **Top right**: image most similar to the query; **middle left**: image ranked second; **middle right**: image ranked third; **bottom left**: image ranked fourth; **bottom right**: last image in the ranked list of five images returned.

6 Length Filtered Orientation Radiograms

In this section we describe an extension to the current ornament image retrieval system. The extension is based on studies of hyper-complex cells in the human visual system. It has been shown that certain cells in the human visual system respond to linear stimuli having not only a specific orientation, but also having a certain range of length. A similar effect can be achieved by filtering the orientation images based on the lengths of the linear structures represented in them. The result is a set of *length-filtered orientation radiograms* for each orientation image. First, the length of each linear structure in an orientation image is estimated. These measurements are used to construct, for each orientation image, several (in our case, three) length-filtered orientation radiograms.

As explained before, each orientation image contains linear structures of the input image that are oriented along the specific pass orientation. We consider the sequential *transverse-* and *longitudinal-projections* of each orientation image. The transverse projection is obtained by projecting the orientation image onto an axis perpendicular (transverse) to its pass orientation. Thus, for an orientation image having the vertical axis (90^o) as the pass orientation, the transverse projection would be obtained by projecting the orientation image onto the horizontal axis. Since the orientation images are gray-valued images, we compute a projection of an orientation image, onto an axis, by summing the gray values of all the pixels that lie on the same line perpendicular to the axis. For example, the projection onto the horizontal axis is computed by summing up the gray values of the pixels that fall on the same vertical line. The peaks in the transverse projection identify the positions of linear structures along the transverse axis. This projection is used to segment the orientation image into several sections, such that each section contains the linear structures represented by *one peak* in the transverse projection. Next, each section of the orientation image is considered in turn. A *longitudinal projection* of *each section* is obtained by projecting the section onto a longitudinal axis, that is, an axis parallel to the pass orientation of the orientation image in question. Each peak in this longitudinal projection corresponds to exactly one linear structure in the section. The length of this linear structure can be computed from the longitudinal projection. In this way, the length of every linear structure in an orientation image can be estimated. Figure 5 describes the projection-

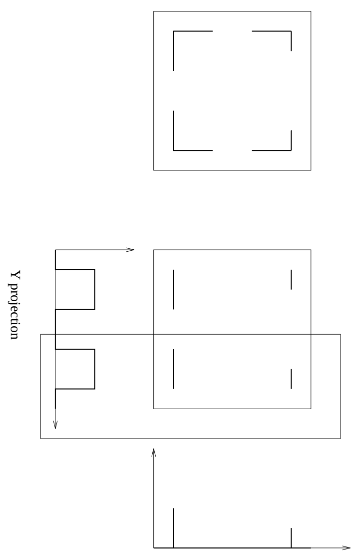

Figure 5: Decomposition of an orientation image using transverse (y) and longitudinal (x) projections. A synthetic image is shown on top. The projection-based decompositions of one orientation image (vertical pass-orientation) are shown below.

based decomposition of orientation images. A synthetic image is shown on top. One orientation image (corresponding to the pass orientation of 90°) and its transverse and longitudinal decompositions are shown below it. In the figure, 'Y projection' shows the transverse projection. Based on this projection, the orientation image can be divided into two transverse sections. 'X projection' shows the longitudinal projection for the lower section. This isolates two linear structures in the example orientation image. The lengths of these regions may also be computed from the projections.

The orientation image is now split into three *length-filtered orientation images*. The linear structures in the orientation image are grouped into three length-ranges, and each group is represented in a separate length-filtered orientation image. Radiograms of the length-filtered orientation images are called *length-filtered orientation radiograms.* Thus, each length-filtered orientation radiogram represents linear structures of the original image that have a specific orientation as well as fall within a specific range of length. Note that this information cannot be derived using simple connected component analysis, since the orientation images are gray-scale images.

Length-filtered orientation radiograms can be used to generate descriptions of *local shape*. We are implementing an image retrieval system using length-filtered orientation radiograms. This idea can be used to process sub-image queries, where the user can request images which contain the query as a part of a larger structure.

7 Conclusion

Existing image retrieval systems that use shape descriptors, rely on a segmentation of an image to estimate the features of the various shapes in the image. Such an approach is inherently limited since the segmentation is static and does not take the user's needs into account. The segmentation provides a very high level of abstraction which limits the utility of the system to a very small set of queries. Moreover, this approach yields boundary-based shape features which are not adequate for characterizing complex shapes, for example the ornament image shown in Figure 1.

In this paper we have described an image retrieval system that utilizes shape information to index images in a collection but does not use a segmentation based approach to derive the shape information. The shape information present in an image is extracted using orientation radiograms. This approach has several advantages. Orientation radiograms represent a lower level of abstraction than a rigid segmentation, and hence are more general in nature. Thus, radiograms can adequately support a larger variety of queries than features derived based on segmentation. Secondly, orientation radiograms are very well suited for characterizing local shape information. One very useful property of orientation radiograms is that they are additive in nature. Thus, an image can be described by an aggregation of the radiograms of its sub-images. This property is can be exploited for retrieval using queries where only a portion of a larger image is specified. Boundary-based global shape measures, such as moments, are not additive in nature.

We use a collection of *ornament* images to present results demonstrating the power of orientation radiograms for image retrieval problems. Currently only queries that specify an entire ornament are supported by the image retrieval system. In this paper we have outlined the extension to our system for handling sub-image queries.

It is true that our current collection pertains to a specific domain of images. An approach based on global histograms of orientations has been successfully used for indexing a collection of images with different textures [7]. Based on this report and our own experiences, we are encouraged to believe that orientation radiograms can indeed be used to characterize shape information in a fairly heterogeneous collection of images. We plan to conduct experiments to determine the utility of orientation radiograms for other kinds of images.

References

[1] M. Flickner et al. Query by Image and Video Content: The QBIC System. *IEEE Computer*, pages 23 – 32, Sept. 1995.

[2] R. Mehrotra and J. E. Gary. Shape-Similar Retrieval in Shape Data Management. *IEEE Computer*, pages 57 – 62, Sept. 1995.

[3] R. W. Picard. Light-Years from Lena: Video And Image Libraries of the Future. In *Proc. Int. Conf. on Image Proc.*, pages 310 – 313, Washington D.C., USA, October 1995.

[4] J. Bigün and G. H. Granlund. Optimal Orientation Detection of Linear Symmetry. In *Proc. First Int. Conf. on Computer Vision*, pages 433 – 438, London, June 1987.

[5] S. Corsini. Vers un Corpus des Ornements Typographiques Lausannois du 18ème Siècle. Problèmes de Définition et de Méthode. In *Ornementation typographique et bibliographie historique, actes du Colloque de Mons*, pages 139–158, August 1988.

[6] S. S. Reddi. Radial and Angular Moment Invariants for Image Identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(2):240 – 242, March 1981.

[7] M. M. Gorkani and R. W. Picard. Texture Orientation For Sorting Photos “At a Glance”. In *Proc. Int. Conf. on Pattern Recognition*, pages 459 – 464, Jerusalem, Israel, September 1994.