CBIR APPROACH TO BUILDING IMAGE RETRIEVAL BASED ON INVARIANT CHARACTERISTICS IN HOUGH DOMAIN

Xiang Yuan and *Chang-Tsun Li*

Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK
{x.yuan, ctli}@dcs.warwick.ac.uk

ABSTRACT

In this paper, we propose two rotation and scale invariant features extracted from the Hough transform domain to guide a CBIR system in the search of relevant building images. Upon receiving a query image, the CBIR system transforms the edges detected from the query into the Hough domain with 180 degrees/bins. From each bin, the *peak percentage* and *peak distance ratio* are calculated. The circular correlations between the peak percentages and peak distance ratios across the 180 bins of the query image and those of the database images are then taken as the similarity measure for ranking the relevance of the database images to the query.

***Index Terms*—**content based image retrieval (CBIR), feature extraction, matching algorithm, Hough transform, image database.

1. INTRODUCTION

Content-based image retrieval has drawn enormous attention in the last decade. As the sizes of image databases grow dramatically, the need for efficient and effective digital image retrieval systems to manage the use of massive information becomes more acute. In CBIR systems, a set of features (feature vector), such as color [1, 2], texture [3, 4] and shape [5-9], is used to represent a digital image. Retrieval is performed by comparing the feature vector of the query image against those of the database images using a matching algorithm. However, due to the fact that the colour and texture of a building image is highly variable according to different illuminations conditions and viewing positions and the fact that human tends to be able to recognize buildings solely by their shapes, shape features have been exploited as a powerful clue for building identification. In [7], the scale invariant feature transform (SIFT) descriptors combined with feature selection mechanism are proposed to detect buildings. In [8], the authors proposed matching descriptors associated with line segmentation to achieve building recognitions. In our previously proposed CBIR system [9], we found that the distribution of Hough peaks in the Hough diagram can represent the distribution of linear signals in building images. To describe the distribution of Hough peaks, the edge map of a building image is extracted first. Secondly, Hough transform is applied to transform the edge map from the spatial domain into the Hough domain. By partitioning the Hough domain into a number of bands, the centroid of the Hough peaks in each band is calculated. Then a band-wise matching (BWM) algorithm is employed to measure the similarity between the query image and the images in the database by taking the centroid set as the feature vector. Experiments showed that the proposed CBIR system is effective in retrieving building images with strong linear features. However, the improperly predefined bandwidth can make the bands of two Hough diagrams miss-matched. Also, significantly rotated relevant building images cannot be retrieved effectively. To overcome the limitations of this CBIR system an improved CBIR system based on a set of rotation and scale invariant features extracted from the Hough transform is proposed in this work.

The rest of the paper is organized as follows. In Section 2, we describe details of the proposed CBIR system. Some experimental results are demonstrated in Section 3. In Section 4, the retrieval performance of the proposed CBIR system is evaluated. Finally, Section 5 concludes the work and points out the future work.

2. PROPOSED CBIR SYSTEM

2.1. Observations

Without using the distance centroid set as we did in [9], an alternative is to calculate the percentage of the Hough peaks distributed in each 1°-wide bin of the Hough diagram. The *peak percentage* in the i^{th} bin is defined as:

$$n_i = n_i / n \tag{1}$$

where n_i is the number of Hough peaks in the *i*th bin, n is the total number of Hough peaks in the whole Hough diagram and is defined as:

$$n = \sum_{i=1}^{180} n_i \tag{2}$$

Figure 1 shows the same building in three different images while Figure 2 shows their corresponding *peak percentage* diagram. From Figure 2, we can see that Figure 2(b) is basically a circularly shifted version of Figure 2(a) while Figure 2(a) and Figure 2(c) have very similar *peak percentage* distributions. Although Figure 2(c) represents the down-scaled image comparing to the image represented by Figure 2(a), the scale problem seems to have no significant impact on the peak percentage values in these peak percentage diagrams.

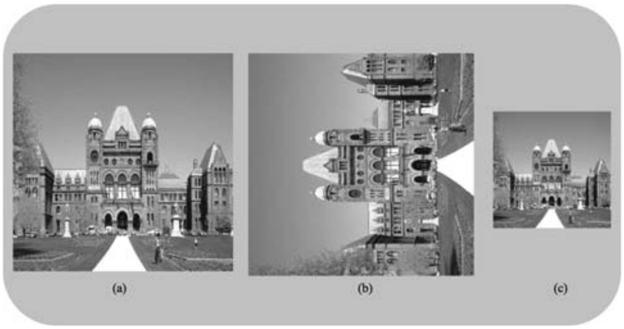

Figure 1. (a) is a grey scale image with the size 512 × 512. (b) is a rotated version of (a). (c) is a down scaled version of (a) with the size 256 × 256.

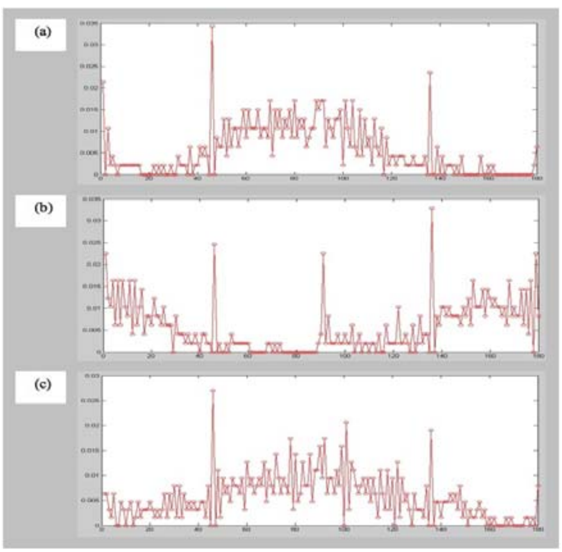

Figure 2. (a), (b) and (c) are the *peak percentage* diagrams of Figure 1(a), (b), and (c), respectively.

We also observed that the ratio of the sum of distances to the centroid of distances (we call it *distance ratio* for short), as defined in Eq. (3) exhibits similar invariant characteristics as peak percentage. The *distance ratio* in the *i*th bin is defined as:

$$r_i = \frac{\sqrt{\sum_{j=1}^{n_i} d_j^2}}{C_i} \tag{3}$$

where d_j is the distance value of the *j*th Hough peaks and the distance centroid C_i of the *i*th bin and is defined as:

$$C_i = \frac{\sum_{j=1}^{n_i} \omega_j \bullet d_j}{\sum_{j=1}^{n_i} \omega_j} \tag{4}$$

where ω_j is the number of points in the edge map that contribute to the formation of the *j*th Hough peak. The *distance ratio* diagrams of Figure 1(a), (b), and (c) are illustrated in Figure 3(a), (b), and (c), respectively.

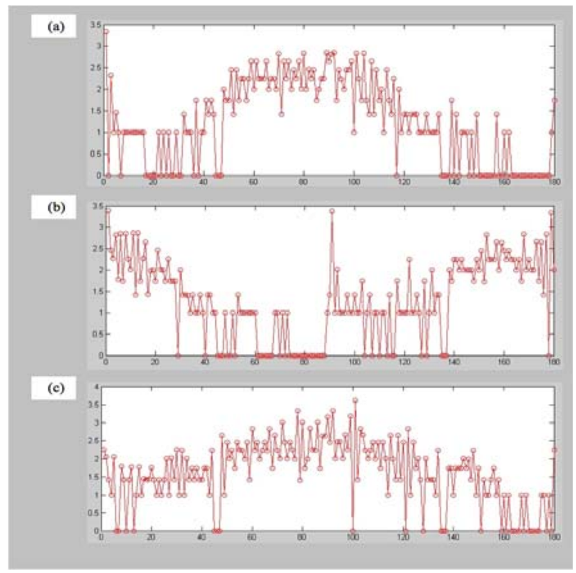

Figure 3. (a), (b) and (c) are the *distance ratio* diagrams of Figure 1(a), (b), and (c), respectively.

2.2. Feature Extraction

Three steps are taken to extract the edge map of a building image. First of all, the image is smoothed with a 3 × 3 Gaussian kernel with the standard derivation equal to 0.5. Secondly, the Canny edge detector is applied to the smoothed image in order to extract the linear information from the image. A thinning process follows the edge detection operation to eliminate the redundant edges. The edge map is then processed by Hough transform. Based on the angle values which are the bin-identifier, all the Hough peaks are classified into 180 bins in the Hough diagram of the edge map. According to Eq.(1) and (2), the peak percentage values for all the bins are extracted. Let $p^q = \{p_1^q, p_2^q, \ldots, p_{180}^q\}$ and $p^k = \{p_1^k, p_2^k, \ldots, p_{180}^k\}$ represent the peak percentage diagrams of the query image and the *k*th database image, respectively. Based on Eq.(3), the distance ratio diagrams of the query image and the *k*th database image can be represented as $r^q = \{r_1^q, r_2^q, \ldots, r_{180}^q\}$ and $r^k = \{r_1^k, r_2^k, \ldots, r_{180}^k\}$, respectively.

2.3. Retrieving Process

In the retrieving process, the circular correlation is employed to measure the similarity between the peak percentage diagram of the query image and that of a database image. It is also the similarity measuring algorithm for the distance ratio diagram of the query image and that of a database image. The *m*th circular correlation between p^q and p^k is defined as:

$$u^k(m) = \frac{\sum_{i=1}^{180-m}(p_i^q \bullet p_{i+m}^k) + \sum_{i=180-m+1}^{180}(p_i^q \bullet p_{i+m-180}^k)}{\sqrt{\sum_{j=1}^{180}(p_j^q)^2} \bullet \sqrt{\sum_{j=1}^{180}(p_j^k)^2}} \tag{5}$$

The maximum circular correlation $\hat{u}^k$ between p^q and p^k is:

$$\hat{u}^k = \arg\max_i \{u^k(i)\} \tag{6}$$

The *m*th circular correlation between r^q and r^k is defined as:

$$v^k(m) = \frac{\sum_{i=1}^{180-m}(r_i^q \bullet r_{i+m}^k) + \sum_{i=180-m+1}^{180}(r_i^q \bullet r_{i+m-180}^k)}{\sqrt{\sum_{j=1}^{180}(r_j^q)^2} \bullet \sqrt{\sum_{j=1}^{180}(r_j^k)^2}} \tag{7}$$

The maximum circular correlation $\hat{v}^k$ between r^q and r^k is:

$$\hat{v}^k = \arg\max_j \{v^k(j)\} \tag{8}$$

For a database with z images, let $\hat{u} = \{\hat{u}^1, \hat{u}^2, ..., \hat{u}^z\}$ and $\hat{v} = \{\hat{v}^1, \hat{v}^2, ..., \hat{v}^z\}$ be sets of circular correlations of the peak percentage diagram and distance ration diagram between the query image and the z database images, and $rank\ (\hat{u}^k)$ and $rank\ (\hat{v}^k)$ be the ranking of $\hat{u}^k$ and $\hat{v}^k$ in $\hat{u}$ and $\hat{v}$, respectively. The final ranking $rank^k$ of the database image, I^k, is given as

$$rank(I^k) = rank(\hat{u}^k) + rank(\hat{v}^k) \tag{9}$$

The database images are then displayed / retrieved according to their rankings, with the one having the highest ranking being displayed first.

3. EXPERIMENTAL RESULTS

In the experiments, the interface is implemented in MATLAB 7.0. There are 235 building images in the database, which are captured by the authors or downloaded from the Google gallery (*http://images.google.com/*). To maintain the diversity of the database, different types of building images are included, e.g., towers, churches, cathedrals, castles, pyramids and general buildings (residential buildings, university departments). The query image is displayed at upper-left corner of the interface and the retrieved images in the right window are displayed according to their rankings in the raster scan (i.e., row first) order.

From Figure 4(a), we can see that 4 relevant images are retrieved by the proposed CBIR system among the 9 top-raking retrieved images while, when given the same query image, as shown in Figure 4(b), our previously proposed system [9] perform relatively poorer, with only 3 relevant images among in the 9 top-ranking images. What is noteworthy is that Figure 4(a) shows that the highly rotated images (the first and fifth retrieved images in the raster scan order), which the previously proposed system [9] cannot handle effectively (see Figure 4(b)), are successfully retrieved by the current system. A broader view of performance of 4 different systems using different shape descriptors are illustrated in Figure 5. The *relevant images – retrieved images* plot in Figure 5 shows that the invariant Hough descriptors proposed in this work outperforms the other three.

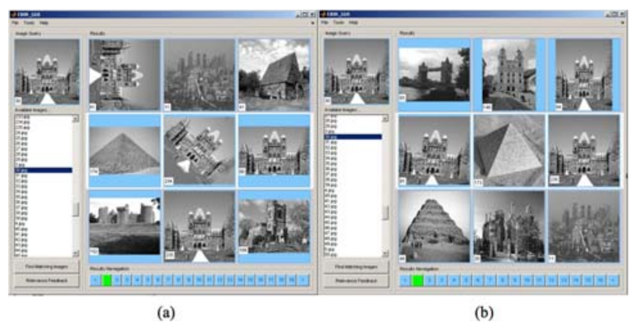

Figure 4. (a) Retrieval results by the proposed invariant Hough descriptors. (b) Retrieval results by previously proposed Hough descriptors.

4. PERFORMANCE EVALUATION

In Figure 2, we observed that the peak percentage diagram of an image is a phase-shifted version of the peak percentage diagram of the rotated version of the same image. The same observation on the distance ratio diagrams is also valid, as demonstrated in Figure 3. The phase needed to shift the peak diagram is very close or equal to the phase needed to shift the distance ratio diagram. As shown in Eq. (6) and Eq. (8), i and j are the phase displacement needed to achieve the maximum circular correlations between the query image and a database image. Therefore, the difference between i and j should be under some threshold λ (see Eq. (10)) if a database image is be deemed as the rotated or scaled versions of the query image.

$$|i-j| \le \lambda \tag{10}$$

The performance of the proposed system with and without enforcing Eq. (10) is illustrated in Figure 6 and Figure 7. We can see the improvement when Eq. (10) is enforced (λ = 5 in this case).

5. CONCLUSION AND FUTURE WORK

In this work, invariant Hough descriptors and the matching algorithm, which employ circular correlation for comparing *peak percentage* and *distance ratio* diagrams, are introduced. From our observations and experiments, we can see that invariant Hough descriptors combined with circular correlation matching algorithm overcome the limitations of our previously proposed CBIR system. We have demonstrated their robustness against rotation and scaling and their superiority over two other shape descriptors. Currently, the robustness of the proposed CBIR system against different viewing aspects is under investigation.

REFERENCES

[1]G. Paa and R. Zabith, "Histogram refinement for content based image retrieval," in *Proc.* WACV, pp. 96-102,1996.

[2]Y. Deng, B.S. Manjunath, C. Kenney, M.S. Moore and H. Shin, "An efficient color representation for image retrieval," *IEEE Trans. On Image Processing,* vol. 10, no. 1, pp.140–147, 2001.

[3]S. Grigorescu, N. Petkov and P. Kruizinga, "Comparision of texture features based on Gabor filters," *IEEE Trans. On Image Processing*, vol. 11, no. 10, pp.1160-1167, 2002.

[4]A.C. Bovic, "Analysis of multichannel narrow band filters for image texture segmentation," *IEEE Trans. Signal Processing*, vol. 39, pp. 2025-2043, 1991.

[5]J.W. Han and L, Guo, "A shape-based image retrieval method using salient edges," *Signal Processing: Image Communication*, vol. 18, no. 2, pp.141-156, 2003.

[6]A. Jain, A. Vailaya and H.J. Zhang, "On image classification: city image vs. landscapes," *Pattern Recognition*, vol. 31, no. 12, pp.1921-1935, 1998.

[7]G. Fritz, C. Seifert, M. Kumar and L. Paletta, "Building detection from mobile imagery using informative SIFT descriptors", in *Proc. SCIA*, pp.629-638, 2005.

[8]T. Geodeme and T. Tuytelaars, "Fast wide baseline matching for visual navigation," in *Proc. CVPR*, pp. 24-29, 2004.

[9]X. Yuan and C.-T. Li, "CBIR approach to building image retrieval based on linear edge distribution", in *Proc. AVSS*, pp. 95-101, 2006.

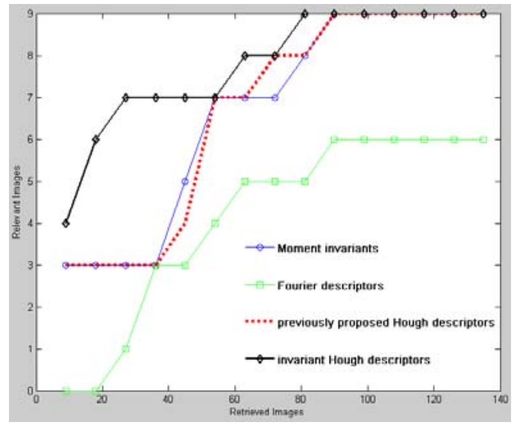

Figure 5. A *Relevant images* - *Retrieved images* plot. The horizontal axis represents the number of retrieved images while the vertical axis represents the number the relevant images among the retrieved images.

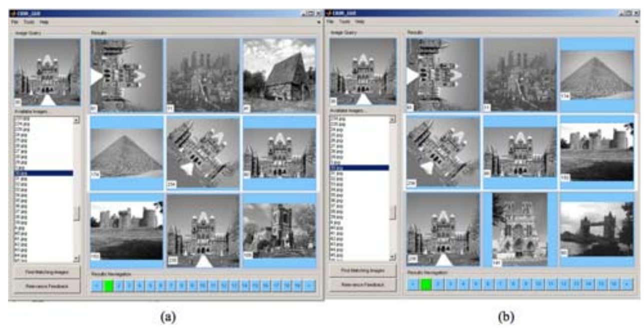

Figure 6. (a) is the retrieved results by invariant Hough descriptors. (b) is the retrieved results by the invariant Hough descriptors after enforcing Eq. (10) where $\lambda = 5$.

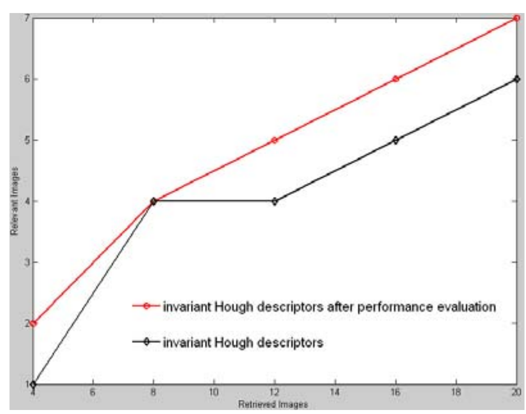

Figure 7 The *Relevant images* - *Retrieved images* plot for invariant Hough descriptors with Eq. (10) enforced ($\lambda = 5$)