CONTENT-BASED IMAGE RETRIEVAL USING GENERIC FOURIER DESCRIPTOR AND GABOR FILTERS

Quan He, ZhengQiao Ji and Q. M. Jonathan Wu
Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada

Keywords: CBIR, Generic Fourier Descriptor, Gabor filters, feature extraction, texture, shape, retrieval.

Abstract: Content-based image retrieval (CBIR) is an important research area with application to large amount image databases and multimedia information. CBIR has three general visual contents, including color, texture and shape. The focus of this paper is on the problem of shape and texture feature extraction and representation for CBIR. We apply Generic Fourier Descriptor (GFD) for shape feature extraction and Gabor Filters (GF) for texture feature extraction, and we successfully combine GFD and GF together for shape and texture feature extraction. Experimental results show that the proposed GFD+GF is robust to all the test databases with best retrieval rate.

1 INTRODUCTION

Due to the emergence of large-scale image collections, Content-Based Image Retrieval (CBIR) was proposed for the need of image content description and representation so that automatic searching is possible (Rui, 2002). Basically, CBIR has three general visual contents: color, texture and shape, and Feature (content) extraction is the basis of content-based image retrieval. The objective of this paper is to study the extraction of both shape and texture features for image retrieval.

Shape is one of the most important visual image features because shape is a very important feature to human perception. Numerous shape descriptors have been proposed in literature, these descriptors can be broadly categorized into two groups: contour-based and region-based descriptors. Fourier descriptor and Zernike moments are the favourite descriptors in these two groups respectively (Zhang, 2004).

Contour-based shape descriptors exploit only boundary information, thus ignoring the shape interior content. Region-based shape descriptors are derived by using all the pixel information within a shape region, so they can be applied to general applications. However, most of region-based shape descriptors are extracted from spatial domain so that they are sensitive to noise and shape variations. In order to overcome the disadvantages, a Generic Fourier Descriptor (GFD) was proposed (Zhang, 2002), GFD is rotation, translation and scale invariant, and experimental results showed that GFD has better performance than the common contour-based and region-based descriptors. So we choose GFD for shape feature extraction in our work.

Texture is also a key component of human visual perception. It contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment (Rui, 2002). Two-dimensional Gabor filters (GF) is proved to be very effective texture feature extraction methods in literature. Ma and Manjunath evaluated the texture images by various wavelet transform representations (Manjunath, 1996). They found that GF had the best performance within the tested candidates. So we choose GF with specified orientations and frequencies for texture features.

In this paper, we consider both shape and texture features for image retrieval. Shape features are extracted by using GFD and texture features are derived by applying GF. The rest of the paper is organized as follows. In section 2, background of GFD and GF are described and the procedure of shape and texture feature extraction is introduced. In section 3, we test the approach on different databases and give the experimental results and discussions. Section 4 concludes the paper.

He Q., Ji Z. and M. Jonathan Wu Q. (2008).
CONTENT-BASED IMAGE RETRIEVAL USING GENERIC FOURIER DESCRIPTOR AND GABOR FILTERS.
In *Proceedings of the Third International Conference on Computer Vision Theory and Applications*, pages 525-528
DOI: 10.5220/0001080105250528

2 FEATURE EXTRACTION

Feature extraction is the basis of content-based image retrieval. Features may include both text-based features (key words, annotations) and visual features (color, texture, shape, faces). We will confine our research to the techniques of shape and texture feature extraction.

2.1 Shape Feature Extraction

In this section, we first describe the shape descriptor GFD in detail. And then we give the procedure of obtaining the feature representation.

2.1.1 Generic Fourier Descriptor (GFD)

Fourier Transform (FT) has been widely used for image processing and analysis. Image in spectral domain is robust to noise and some kind of image distortions. However, Applying 1-D FT to shape indexing involves the knowledge of the shape boundary, while some image boundary may not be available. 2-D FT can be directly applied to any shape image and thus overcome this limitation.

Here are formulas for the continuous and discrete 2-D Fourier transform of a shape image $f(x, y)$ $(0 \le x < M, 0 \le y < N)$.

$$F(u,v) = \int_x \int_y f(x,y) \times \exp[-j2\pi(ux+vy)]dxdy \quad (1)$$

$$F(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y) \times \exp[-j2\pi(ux/M+vy/N)] \quad (2)$$

However, direct applying the 2-D FT to the Cartesian representation of an image is not practical. Because the resulting descriptors are not rotation invariant, which is a crucial property for a shape descriptor. Zhang et al. give a solution to this problem by applying 2-D FT on polar-raster sampled shape image (Zhang, 2002). Furthermore, In order to find a correspondence between input parameters and their physical meaning, Zhang et al. proposed to treat the polar image in polar space as a normal two-dimensional rectangular image in Cartesian space, as shown in Figure 1.

$$PF(\rho,\varphi) = \sum_r \sum_i f(r,\theta_i) \times \exp[-j2\pi(\frac{r}{R}\rho + \frac{2\pi i}{T}\varphi)] \quad (3)$$

Where $0 \le r < R$, $\theta_i = i(2\pi/T)$ $(0 \le i < T)$; $0 \le \rho < R$ and $0 \le \varphi < T$. R and T are the radial and angular resolutions. $f(x, y)$ is a binary function in shape application.

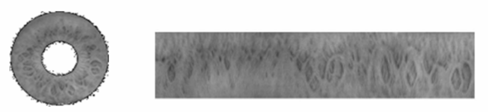

Figure 1: Polar representation of an image

2.1.2 Shape Feature Representation

In order to obtain invariant GFD features, suitable normalization has been applied. Translation invariant is achieved by setting the centroid of the shape to be the origin of the polar space. The Polar Fourier Transform (PFT) operator is then applied on the normalized image. In order to achieve scale invariance, the magnitude values are normalized by the magnitude of the first coefficient, and the first magnitude value is normalized by the area of the circle. Rotation invariance is achieved by ignoring the phase information in the coefficients. By selecting n radial frequencies and m angular frequencies, the resulting real values are organized in a linear feature vector as follows:

$$FD = \left\{ \frac{|PF(0,0)|}{area}, \ldots, \frac{|PF(0,n)|}{|PF(0,0)|}, \ldots, \frac{|PF(m,0)|}{|PF(0,0)|}, \frac{|PF(m,n)|}{|PF(0,0)|} \right\} \quad (4)$$

This feature vector is normalized to range [0, 1] by normalization as follows:

$$FD = \frac{FD - \min(FD)}{\max(FD) - \min(FD)} \quad (5)$$

For two shapes represented by their GFD, the similarity between the two shapes is measured by the Euclidean distance. In our implementation, we use 36 GFDs (3 radial frequencies and 12 angular frequencies) and 60 GFDs (4 radial frequencies and 15 angular frequencies) indicated in literature (Zhang, 2002) for feature extraction.

2.2 Texture Feature Representation

In this section, we describe texture representation based on Gabor filters, and also the normalization of features.

2.2.1 Gabor Filters (GF)

Gabor filters are a group of wavelets, with each wavelet capturing energy at a specific frequency and direction. Typically, an input image $I(x, y)$ with size $P \times Q$, is convolved with a 2D Gabor function $g_{mn}(x, y)$, to obtain a Gabor feature $G_{mn}(x, y)$ as follows:

$$G_{mn}(x,y) = \sum_{x_1}\sum_{y_1} I(x_1,y_1) g_{mn}^{*}(x-x_1, y-y_1) \quad (6)$$

Where * indicates the complex conjugate.

A 2D Gabor function $g(x,y)$ has its form:

$$g(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp[-\frac{1}{2}(\frac{x^2}{\sigma_x}+\frac{y^2}{\sigma_y})+2\pi jWx] \quad (7)$$

And $g_{mn}(x,y)$ is a set of self-similar functions generated from dilation and rotation of the Gabor function $g(x,y)$ (Manjunath, 1996).

$$g_{mn}(x,y)=a^{-m}g(x',y') \quad (8)$$

$$x'=a^{-m}(x\cos\theta+y\sin\theta),$$

$$y'=a^{-m}(-x\sin\theta+y\cos\theta) \quad (9)$$

Where $m=0,1,\dots M-1$, $n=0,1,\dots N-1$, m and n specify the number of scales and orientations respectively, and $a>1$, $\theta=n\pi/N$.

2.2.2 Texture Feature Representation

We can obtain a set of magnitudes by applying Gabor filters on the image $I(x,y)$ with different orientation at different scale.

$$E(m,n)=\sum_x\sum_y|G_{mn}(x,y)| \quad (10)$$

The mean μ_{mn} and standard deviation σ_{mn} of the magnitude of the transformed coefficients are as follows:

$$\mu_{mn}=\frac{E(m,n)}{P\times Q},$$

$$\sigma_{mn}=\frac{\sqrt{\sum_x\sum_y(|G_{mn}(x,y)|-\mu_{mn})^2}}{P\times Q} \quad (11)$$

The Gabor feature vector is given by:

$$f=[\sigma_{00},\sigma_{01},\dots,\sigma_{(M-1)(N-1)}] \quad (12)$$

And then we normalize the features to [0, 1] as described in (5). The similarity between two texture features is measured by the Euclidean distance. In our implementation, five scales and six orientations are used for feature extraction.

2.3 Combined Shape and Texture Features

Based on Generic Fourier Descriptor and Gabor filters, we obtain GFD&GF feature vector for image retrieval. The overall distance between the query image I_q and the database image I_d is as follows:

$$D(I_q,I_d)=w_{GFD}D_{GFD}(I_q,I_d)+w_{GF}D_{GF}(I_q,I_d) \quad (13)$$

One way of choosing w_{GFD} and w_{GF} is to use relevance feedback for iterative setting (Lee 2002). The problem of relevance feedback is outside the scope of paper, we use $w_{GFD}=w_{GF}=0.5$ in our experiments.

3 EXPERIMENTAL RESULTS

We have conducted performance tests both on shape images and texture images, a set of comparison for GFD and GF are made in this section. The retrieval rate for the query image is measured by counting the number of images from the same category that are found in the top m matches (Bimbo 1999).

3.1 Experiment 1 - Shape Database

The silhouette database is collected by The Laboratory for Engineering Man/Machine Systems (LEMS) in Brown University. Our silhouette database consists of 600 shape images with 30 subjects and 20 images per subject. Each image in the database is indexed using GFD, GF and combined features, 30 images (one image per class) are randomly selected as queries.

From table 1, we can see that GFD and GFD+GF are similar in retrieval rate and 36GFDs+GF gives the highest retrieval rate. GFD with 3 radial and 12 angular resolutions has a little higher performance than GFD with 4 radial and 15 angular resolutions. GF has the lowest performance. By combining GF with GFD, we can obtain better retrieval rate. So this experiment shows that the combined features are effective for shape-base image retrieval with best performance.

Table 1: Retrieval Rate on LEMS silhouette database.

	Number of top matches				
Methods	1	5	10	15	20
GF	100	84.04	53.24	47.83	43.09
36GFDs	100	95.62	82.05	68.37	43.69
60GFDs	100	97.97	78.53	66.89	42.83
36GFDs+GF	**100**	**96.57**	**82.5**	**69.67**	**51.86**
60GFDs+GF	100	95.5	80.5	64.45	44.81

3.2 Experiment 2 - Fingerprint Database

This database is collected by FVC2000. It contains 110 fingers wide (w) and 8 impressions per finger.

50 images from different finger categories are randomly selected as queries.

Table 2: Retrieval Rate on LEMS Fingerprint database.

	Number of top matches				
Methods	1	2	4	6	8
GF	100	100	95.36	82.25	39.23
36GFDs	100	91.23	57.23	31.07	22.22
60GFDs	100	91.62	58.01	31.77	22.27
36GFDs+GF	100	100	93.9	77.98	51.12
60GFDs+GF	**100**	**100**	**96.54**	**82.72**	**54.24**

It shows from table 2 that 60GFDs+GF achieves the best retrieval rate. GF gets very high retrieval rate for fingerprint database. There is little difference (2%) between GF and GFD+GF in the retrieval rate, but GFD gives very low retrieval rate. The results indicate that GF and GFD+GF are much more effective for fingerprint database than GFD.

3.3 Experiment 3-Object Image Database

This database is collected by Amsterdam Library of Object Images (ALOI). ALOI is a color image collection of one thousand small objects. The images are systematically varied from viewing angle, illumination angle, and illumination color for each object.

Our database consists 1200 images, and it's organized into 50 groups while 24 similar images in each group. In our experiment, we use gray level image for observation since we only extract shape and texture features of images. 50 images (one image per class) are randomly selected as queries.

Table 3: Retrieval Rate on ALOI database.

	Number of top matches				
Methods	1	4	12	20	24
GF	100	89.09	67.84	55.18	36.05
36GFDs	100	100	73.63	54.3	38.11
60GFDs	100	100	87.6	57.5	42.8
36GFDs+GF	100	100	88.64	76.52	49.77
60GFDs+GF	**100**	**100**	**89.02**	**78.10**	**59.55**

It can be seen from Table 3 that 60GFDs+GF outperforms GFD and GF on ALOI database. Both of 36GFDs and 60GFDs achieve high retrieval rate, while GF has the lowest performance. In table 3, it shows 60GFDs has higher performance (average 5%) than 36GFDs. Although GF has the lowest retrieval rate, the overall performance is still good and it significantly improves the retrieval rate (average 13%) when combine GF and GFD together for image retrieval. The results indicate that both GFD and GF are effective for ALOI database and the combination of GFD and GF achieve the highest retrieval rate and effectively improve the overall performance.

4 CONCLUSIONS

In this paper, we investigate and compare shape feature extraction by GFD and texture feature extraction by GF on three different databases. From the experiment results, we come to a conclusion that GFD+GF can be used as a robust feature of shape and texture for image retrieval with highest performance. GFD is effect for shape database and ALOI database while it's quite weak for fingerprint database, since GFD is a shape descriptor. GF gets very high retrieval rate for fingerprint database, but gives relatively low performance for shape database and ALOI database, since GF mainly extracts texture feature. Scale and translation invariance are to be considered for GF and the selection of weight factors is left for future work.

REFERENCES

Bimbo A.D., 1999. *Visual Information Retrieval*, Morgan Kaufmann Publishers, Inc. San Francisco, USA.

Lee Kyoung-Mi, Nick W., 2002. Street: Incremental feature weight learning and its application to a shape-based query system, *Pattern Recognition Letters,* v.23 n.7. pp 865-874.

Manjunath B.S. and Ma W.Y., 1996. Texture features for browsing and retrieval of image data, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol.18, no.8, pp 837-42.

Rui Y., Huang T.S., and Chang S.F., 1999. Image Retrieval: Current Techniques, Promising Directions, and Open Issues, *Journal of Visual Communication and Image Representation*, Volume 10, pp 39-62.

Zhang D.S. and Lu G., 2004. Review of shape representation and description techniques, *Pattern Recognition*, Vol. 37, No. 1, pp 1-19.

Zhang D.S. and Lu G., 2002. Shape-based image retrieval using generic Fourier descriptor, *SP:IC(17)*, No. 10, pp 825-848.

LEMS Vision Group, http://www.lems.brown.edu/vision/software/index.html

FVC2000, http://bias.csr.unibo.it/fvc2000/

Amsterdam Library of Object Images (ALOI), http://staff.science.uva.nl/~aloi