An Effective Image Retrieval System using Region and Contour based Features

Chandan Singh
Professor
Department of Computer Science,
Punjabi University,
Patiala-147002, Punjab, India.

Pooja
Research Scholar
Department of Computer Science,
Punjabi University,
Patiala-147002, Punjab, India.

ABSTRACT
In this paper, a hybrid approach is proposed for improving the image retrieval accuracy. In the hybrid approach both local and global features of images are combined, which represent the entire aspects of images. Local features are extracted using Fourier descriptors and global features are extracted by means of angular radial transform. The results of combining both these descriptors demonstrate that the proposed solution provides significant improvement in image retrieval accuracy as compared to using global and local features autonomously. Besides, it also outperforms existing recent hybrid approaches to image retrieval.

General Terms
Digital image processing, angular radial transform

Keywords
Angular radial transform, Fourier descriptor, Zernike moments.

1. INTRODUCTION
Digital images are one of the effective media to present a wide variety of information graphically. With the advancement in technology, a huge amount of images are being generated everyday in medical diagnosis, military and defense services, registration purposes, geographical images, trademarks, chromosomes, etc. Therefore, the indexing and retrieval of such visual content has become an active research area [1,2]. Several techniques are developed to index and retrieve images accurately and efficiently. Basically, image can be described by three low level visual features such as color, texture, and shape. Among them, shape feature is considered to be the most effective due to its consistency with human visual perception. Shape is all the geometric information that remains intact when location, scale, and rotation effects are filtered out from the object. As per Kim and Kim [3], a good shape representation should be compact and retain the essential characteristics of a shape.

The existing shape description techniques generally fall into two categories: contour based and region based. The contour based methods such as Fourier descriptor (FD) [4], curvature scale space (CSS) [5], elastic matching [6], chain codes [7], etc, extract information only from the boundary of the shape and neglect the essential information contained in the shape interior. Contour based descriptors are prone to image noise because they consider small amount of image information, i.e., contour only. On the other side, region based methods such as Zernike moments descriptor (ZMD) [3], angular radial transform (ART) [8], moment invariants (MI) [9], generic Fourier descriptor (GFD) [10], etc, consider the entire information of an image. Therefore, these descriptors are more robust to noise and other variations in the image. Hu's seven MI are derived from the low order geometric moments, which are insufficient to describe the complete aspects of an image. Besides, higher order moments are complex to derive and suffer from numerical instability. In GFD, the unstable log polar mapping makes it prone to noise and other transformations. ZMD provides good retrieval accuracy, but they are computation intensive due to high degree of polynomials and a large number of factorial terms in their radial kernel function. ART is adopted by MPEG-7 [8] as a region based shape descriptor, which exhibits similar properties as that of ZMD. Moreover, ART is quite efficient and provides a large number of features, which are numerically stable.

Although contour based descriptors provide the local characteristics of an image and region based descriptors provide the global ones, both local and global features are desirable to describe all the attributes of an image because global and local features are complimentary to each other. Therefore, in a recent work, Wei et al. [11] propose a hybrid approach to trademark image retrieval in which local features are obtained by computing the mean of curvature at each boundary point and mean and standard deviation of distance between centroid and each edge point in edge map. The features based on curvature values have poor performance because the adjacent points do not yield satisfactory curvature of the edges represented by the edge points. In their approach, for global feature representation only four low orders of Zernike moments are used. We observe that moments up to first four orders of Zernike moments are not adequate for describing the global features of an image. Another hybrid approach to image retrieval is proposed by Qi et al. [12] in which local features are extracted by using the relationship among two adjacent boundary points and the centroid (RAPC) and histograms of centroid distances (HCD), representing contour based shape descriptor. Global features are extracted using spatial distribution of feature points and feature point matching. Although HCD provides satisfactory performance, the performance of RAPC is not satisfactory. The radius of the circumscribed circle passing through two close boundary

points and the centroid becomes quite large, which results in incorrect formation of histogram that is unable to distinguish features of one image from another. A survey of local and global feature descriptors [13] concludes that FD and CSS are among the best contour based shape descriptors, while ZMD and ART are observed to be the best among region based descriptors. A recent comparative study on FD and CSS finds FD to be better [4]. Similarly, ART is found to be the best region based descriptor, when a good tradeoff between retrieval performance and computation time is required [14]. Hence, in this paper, we consider both local and global features. For describing the local features, we use FD because it provides accurate association among adjacent boundary points with appropriate number of features. For global features, we use ART because it is quite efficient and adopted by MPEG-7 as a region based shape descriptor. Moreover, magnitude of FD and ART possess rotation invariance property. Then both local and global features are combined to achieve higher retrieval rate as compared to that attained by using these prominent descriptors autonomously. Besides, the proposed system overpowers the performance of recent hybrid approaches TCS and QLS (we refer the hybrid method proposed by [11] as TCS and that of [12] as QLS in the rest of the paper). The performance of the system is evaluated over two standard image databases: MPEG-7 CE shape-1 part B and Kimia-99, which contain images with large intra class variation, distorted, and partial occluded shapes.

2. ANGULAR RADIAL TRANSFORM (ART)

ART is a complex orthogonal unitary transform defined on a unit disc based on complex orthogonal sinusoidal basis functions in polar coordinates. Mathematically, ART is defined as [8]:

$$A_{pq}=\frac{1}{2\pi}\int_0^{2\pi}\int_0^1 f(\rho,\theta)\, V^*_{pq}(\rho,\theta)\rho\, d\rho\, d\,\theta, \qquad (1)$$

where $f(\rho,\theta)$ is an image intensity function in polar co-ordinates, and $V^*_{pq}(\rho,\theta)$ is the kernel function of ART, which is the complex conjugate of $V_{pq}(\rho,\theta)$ defined by

$$V_{pq}(\rho,\theta)=R_p(\rho)e^{jq\theta}, \qquad (2)$$

$$R_p(\rho)=\begin{cases}1 & p=0\\ 2\cos(\pi p\rho) & p>0\end{cases}, \qquad (3)$$

where $p=0,1,2,...,|q|=0,1,2,...,$ are the order and repetition of ART, respectively. We use ART coefficients with orders $(p<3,\ q<12)$, i.e., 35 coefficients to maintain consistency with MPEG-7 standards. For establishing the rotation invariance, only magnitude components of ART are used. For scale invariance, the coefficients of ART are divided by the magnitude of the first coefficient, i.e., $|A_{00}|$. The translation invariance is achieved by considering centre of mass as the centre of the image.

Although ART is computationally very fast, we make it faster by using recursive relations for the calculation of trigonometric functions involved both in radial and angular parts of the kernel function $V^*_{pq}(\rho,\theta)$ and their 8-way symmetry/anti- symmetry [16]. We refer it as ART (recursive) in the rest of the paper.

3. ZERNIKE MOMENTS DESCRIPTOR (ZMD)

The set of orthogonal Zernike moments for an image intensity function $f(\rho,\theta)$ with order p and repetition q are defined over a continuous unit disk $0\le\rho\le1,\ 0\le\theta<2\pi$ [3] as:

$$Z_{pq}=\frac{p+1}{\pi}\int_0^{2\pi}\int_0^1 f(\rho,\theta)V^*_{pq}(\rho,\theta)\,\rho\, d\rho\, d\,\theta \qquad (4)$$

where $V^*_{pq}(\rho,\theta)$ is the complex conjugate of the Zernike polynomials $V_{pq}(\rho,\theta)$, defined as:

$$V_{pq}(\rho,\theta)=R_{pq}(\rho)e^{jq\theta}, \qquad (5)$$

where

$$p\ge0,\ 0\le|q|\le p,\ p-|q|=even,\ j=\sqrt{-1},\ \ and\ \theta=\tan^{-1}(y/x).$$

The radial polynomials $R_{pq}(\rho)$ are defined by:

$$R_{pq}(\rho)=\sum_{k=0}^{(p-|q|)/2}(-1)^k\frac{(p-k)!}{k!\left(\frac{p+|q|}{2}-k\right)!\left(\frac{p-|q|}{2}-k\right)!}\rho^{p-2k} \qquad (6)$$

In our experiments, we use moment order $p_{\max}=12$ as an appropriate order for feature selection. The moment order $p_{\max}=12$ is considered to be a good tradeoff between the computation complexity and image description capability [15]. As in the case of ART, we use fast methods for the computation of ZMD also [16].

4. FOURIER DESCRIPTOR (FD)

FD is obtained by applying Fourier transform on a shape signature. In the first step of the derivation of FD, the boundary coordinates $(x(u),y(u));u=0,1,2,...,N-1$ are obtained, where N represents total number of boundary points. For establishing the translation invariance, radial distance between the boundary points $(x(u),y(u))$ and the centroid (x_c,y_c) of the shape is represented as [4]:

$$r(u)=\sqrt{(x(u)-x_c)^2+(y(u)-y_c)^2} \qquad (7)$$

where $x_c=\frac{1}{N}\sum_{u=0}^{N-1}x(u),\ y_c=\frac{1}{N}\sum_{u=0}^{N-1}y(u)$

The discrete Fourier transform of $r(u)$ is given as:

$$a_n = \frac{1}{N}\sum_{u=0}^{N-1} r(u)\,\exp\left(\frac{-j2\pi u}{N}\right), n = 0,1,..., N-1. \quad (8)$$

The coefficients a_n are called the Fourier descriptors of the shape, and denoted as FD_n. The rotation invariance is achieved by considering only the magnitude and scale invariance is done as follows [4]:

$$F = \left[\frac{|FD_1|}{|FD_0|}, \frac{|FD_2|}{|FD_0|}, ..., \frac{|FD_{N/2}|}{|FD_0|}\right] \quad (9)$$

5. EXPERIMENTAL STUDY AND PERFORMANCE EVALUATION

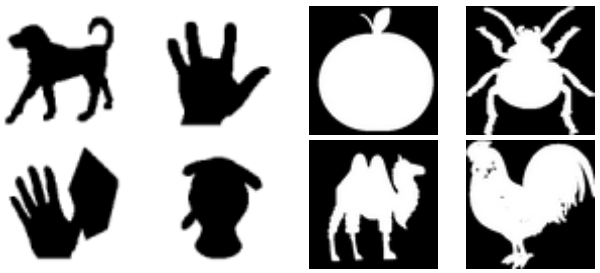
The acquisition of features is done by using the above described methods and their respective features are matched by using the Euclidean distance, which is defined as follows:

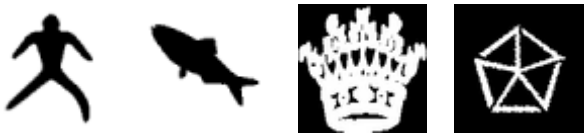
$$ED(Q, D) = \sqrt{\sum_{i=1}^{L}\left(F_i^Q - F_i^D\right)^2} \quad (10)$$

where L is the total number of features of the respective method and F_i^Q and F_i^D represent the feature vectors of the query image and the database image, respectively. The performance of the three state of the art methods ART, ZMD, and FD is evaluated individually and then by combining the local features obtained by FD with the global features obtained by ART and ZMD, i.e., ART+FD and ZMD+FD. In addition, the proposed system is evaluated against the most recent two hybrid approaches TCS [11] and QLS [12]. All algorithms are implemented in VC++ in Microsoft Windows Environment on a PC with 3 GB RAM and Core 2 Duo processor. The image retrieval performance is evaluated over two standard databases, which are described as follows:

- Kimia-99: It contains 9 classes of binary images with 11 instances in each class. The 9 classes consist of planes, fishes, hands, rabbits, etc. The variations include distortions, partial occlusion by other objects, different poses, styles, etc.
- MPEG-7 CE shape-1 part B: It consists of 1400 images containing 70 classes of images with 20 instances in each class. This database represents significant variations within instances of a class.

Some of the instances of database images are given in Figure 1.





Kimia-99 **MPEG-7**

Fig 1: Some of the instances of Kimia-99 and MPEG-7 databases.

The system retrieval accuracy is measured in terms of precision (P) and recall (R) rates. A precision rate can be defined as the percent of retrieved images n similar to the query image among the total number of retrieved images T, i.e., $P = n/T$. A recall rate is defined as the percent of retrieved images n, which are similar to the query among the total number of images D similar to the query in the database, i.e, $R = n/D$. It can be easily seen that both precision and recall rates are functions of the total number of retrieved images. For high retrieval accuracy, the system needs to have both high precision and high recall rates.

5.1 Image retrieval results

As described earlier the retrieval performance of the system is measured in terms of precision and recall. Therefore, here we present the $P-R$ curves for all three methods ART, ZMD, FD autonomously and the proposed ART+FD and ZMD+FD. The $P-R$ curves for Kimia-99 database are given in Figure 2, which depicts that the retrieval effectiveness of QLS is lowest among other methods, followed by TCS, whereas the performance of ART and ZMD is superior to that of FD. By combining the local features with global features, i.e., ZMD+FD, we see that the performance does not improve and its $P-R$ curve coincides with that of ZMD. On the other side, by combining FD with ART, we observe that the retrieval performance significantly improves, this impression provides an important observation that global features based on ART and local features based on FD are more complimentary to each other as compared to ZMD and FD. Another evaluation is performed over MPEG-7 database, and the $P-R$ curves are presented in Figure 3, which reveals that the performance of QLS is still worst than other methods. However, TCS perform better than FD. ART and ZMD perform similar and their $P-R$ curves overlap with each other. However, by combining ZMD with FD and ART with FD, it is observed that the retrieval performance effectively improves and the combination of ART+FD is still superior to that of ZMD+FD.

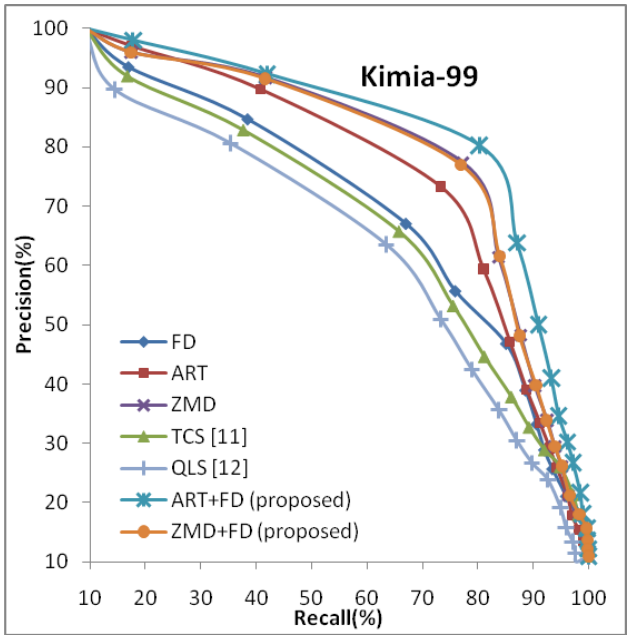


Fig 2: $P-R$ performance of existing prominent descriptors and the proposed hybrid descriptors for Kimia-99 database.

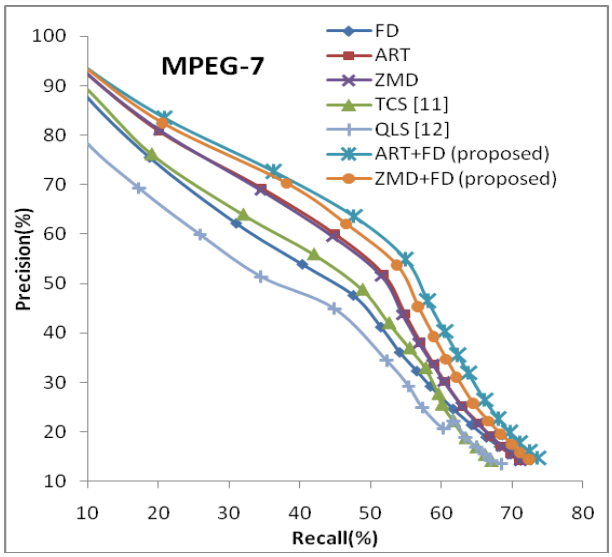


Fig 3: $P-R$ performance of existing prominent descriptors and the proposed hybrid descriptors for MPEG-7 database.

5.2 Robustness to noise

In order to evaluate the robustness of the proposed methods for noise affected images, we consider a Noise database. This database is built by choosing 70 images from each class of MPEG-7 database and Gaussian noise is added by setting mean $\mu = 0$ and standard deviation σ varying between 0.0 and 0.19 with an increment of 0.01, creating a database of 1400 images, where the image with $\mu = 0.0, \sigma = 0.0$ represents the original image. The $P-R$ curves for this noise test are depicted in Figure 4. It is observed that FD is unable to recognize noise affected images. The performance of ZMD is superior to that of ART. By combining FD with ART and ZMD, no improvement is observed in their performance and the $P-R$ curves of ART+FD and ZMD+FD overlap with autonomous ART and ZMD, respectively. The performance of TCS and QLS is also good but not more than ART and ZMD.

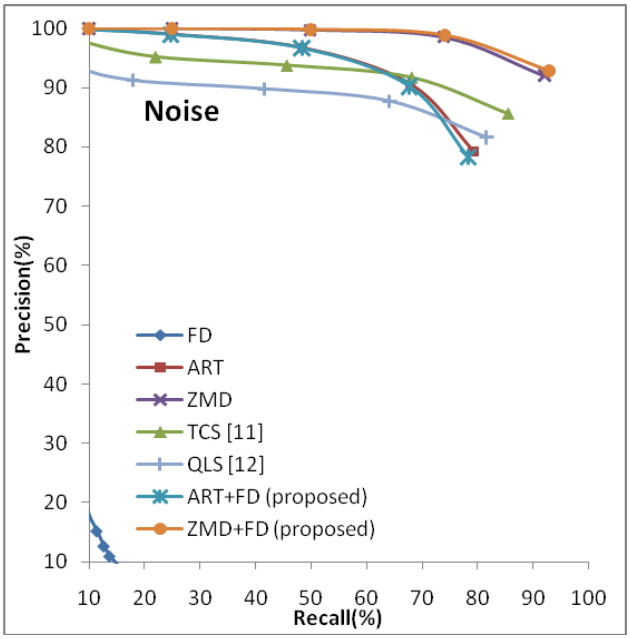


Fig 4: $P-R$ performance of existing prominent descriptors and the proposed hybrid descriptors for Noise database.

5.3 Top retrieval results

The top 11 retrieval results for Kimia-99 database by the proposed and other methods are presented in Figure 5. It is apparent from the figure that the proposed ART+FD is capable of retrieving all relevant images to the query image. The worst performance is given by QLS. The irrelevant images are shaded in grey.

5.4 Computation time analysis

The computation complexity of ART, ZMD, FD are presented in Table 1. The minimum time is taken by FD and the maximum CPU time is elapsed by ZMD. Nevertheless, by computing ZMD and ART using the recursive methods [16], the computation time considerably reduces. The least amount of CPU elapsed time is taken by ART (recursive) followed by FD. Therefore, the proposed hybrid method involving ART+FD is quite efficient by taking only 52.37s on average over a database of 1400 images for computing the features in offline mode. On the other hand, ZMD+FD takes 91.85s for offline feature extraction and 0.091s for online feature extraction. Therefore, ART+FD are quite efficient along with its significant image retrieval accuracy.

6. CONCLUSION

In this paper, the major region based descriptors ART and ZMD and major contour based descriptor FD are evaluated against each other. Besides, a hybrid approach is proposed which combines both local features extracted by FD and global features extracted by ART and ZMD. From the experimental results of $P-R$ curves and CPU elapsed time, it is observed that the proposed hybrid method ART+FD outperforms the traditional autonomous ART, ZMD, and FD in terms of both retrieval rate and CPU elapsed time. The performance of hybrid ZMD+FD also improves but still slightly poorer than ART+FD. The performance of other hybrid approaches TCS and QLS is poor because in TCS inadequate number of global and local features is used. In QLS the local features RAPC leads to inaccurate formation of histograms by which their retrieval performance significantly reduced. Hence, the proposed system based on effective contour and region based descriptors is quite effective and efficient and suitable for robust image retrieval.

7. ACKNOWLEDGMENTS

We are thankful to the University Grant Commission (UGC), New Delhi, India, for providing research fellowship to the second author for carrying out the research work leading to PhD degree.

Table 1. Average CPU elapsed time in seconds for offline and online feature extraction using ART, ZMD, and FD.

Feature extraction	ART	ART (recursive)	ZMD	ZMD (recursive)	FD	ART+FD (recursive)	ZMD+FD (recursive)
Offline (1400 images)	289.22	21.02	375.3	60.27	29.68	52.37	91.85
Online (1 image)	0.109	0.016	0.268	0.064	0.021	0.04	0.091

8. REFERENCES

[1] Veltkamp, R., Burkhardt, H., Kriegel, H.-P., 2008. State-of-the-Art in content-based image and video retrieval, Kluwer Academic Publishers, New York.

[2] Long, F., Zhang, H.J., Feng, D.D., 2003. Fundamentals of content-based image retrieval, in: D.D. Feng, W.C. Siu, H.J. Zhang (Eds.), Multimedia Information Retrieval and Management—Technological Fundamentals and Applications, Springer 1–27.

[3] Kim, W.-Y., Kim, Y.-S., 2000. A region based shape descriptor using Zernike moments, Sig Process: Image Commun. (16) 95–102.

[4] Zhang, D., Lu G., (2003) A comparative study of curvature scale space and Fourier descriptors for shape based image retrieval, J. Vis. Commun. Image R. (14) 41-60.

[5] Mokhtarian, F., Abbasi, F., Kittler, J., 1997. Efficient and robust retrieval by shape content through curvature scale space, Int. workshop on Image Databases and Multi-Media Search 51–58.

[6] Attalla, E., Siy, P., 2005. Robust shape similarity retrieval based on contour segmentation polygonal multiresolution and elastic matching, Pattern Recoginit. (38) 2229–2241.

[7] Freeman, H., 1961. On the encoding of arbitrary geometric configurations. IRE Trans. Electron. Comput. (10) 260–268.

[8] Bober, M., 2001 MPEG-7 visual shape descriptors, IEEE Trans. Circuits Syst. Video Technol. (11) 716-719.Hu, M.-K., 1962. Visual pattern recognition by moment invariants, IRE Trans. Inf. Theory 8 (2) 179–187.

[9] Zhang, D., Lu, G., 2002. Shape-based image retrieval using generic Fourier descriptor, Signal Process.: Image Commun. 17 (10) 825–848.

[10] Wei, C.H., Li, Y., Chau, W.-Y., and Li, C.-T., 2008. Trademark image retrieval using synthetic features for describing global shape and interior structure, Pattern Recognit. (42) 386–394.

[11] Qi, H., Li, K., Shen, Y., Qu, W., 2010. An effective solution for trademark image retrieval by combining shape description and feature matching, Pattern Recognit. (43) 2017–2027.

[12] D. Zhang, G. Lu, Evaluation of MPEG-7 shape descriptors against other shape descriptors, Multimed. Syst. 9 (2003) 15-30.

[13] A. Amanatiadis, V.G. Kaburlasos, A. Gasteratos, S.E. Papadakis, Evaluation of shape descriptors for shape based image retrieval, IET Image Process. 5 (5) (2011) 493-499.

[14] Singh, C., Pooja, 2011. Improving image retrieval using combined features of Hough transform and Zernike moments, Opt. Laser Eng. (49) 1386-1396.

[15] Singh, C., Walia, E., 2011. Algorithms for fast computation of Zernike moments and their numerical stability, Image Vis. Comput. 29 (4) 251-25

(a)
(b)

Methods	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7	Rank 8	Rank 9	Rank 10	Rank
FD											
ART											
ZMD											
TCS											
QLS											
RT+FD											
MD+FD											

Fig. 5(a) Query image (b) Top 11 retrieval results by the proposed and other methods for Kimia-99 database.