A Zernike Moment based Modified CBIR System with Canny Edge Detector

Kushik Bharadwaj
Assistant Professor,
Department of Electronics & Comm. Engg.,
JIET SETG, Jodhpur, Rajasthan, India

Gaur Sanjay B.C.
Associate Professor,
Department of Electronics & Comm. Engg.,
JIET SETG, Jodhpur, Rajasthan, India

ABSTRACT
The moment invariants are used as a feature space for pattern recognition. Shape and texture representation is a fundamental issue in the newly emerging multimedia applications. This work addresses the problem of retrieval in case of rotation, shape, resizing, and translation of an image in the content based image retrieval system. A modified Zernike moment with canny edge detector can be used for the retrieval of an image for such cases. The proposed method is very useful and efficient to retrieve a query image from a very huge complex database. The method has been tested over 400 images from four different groups like Cars (automobiles), Vegetables, Human faces, Natural images etc. Through the result it is found that the proposed method works equally better in all kind of images than the available techniques.

Keywords
Content Based Image Retrieval, Invariant Features, Zernike moments, Canny Edge Detector, Euclidian Distance.

1. INTRODUCTION
In Content Based Image Retrieval System (CBIR), shape and texture are the key visual features to human for distinguishing visual data along with other features of texture and color. In the context of content based image retrieval applications, the main focus of a shape descriptor is measuring perceptual similarity of different shapes. A shape descriptor is required to be invariant to rotation, translation and scale changes. In case of single contour based shapes, high curvature points are usually regarded as feature points and used for obtaining shape descriptors such as tangent space method, curvature scale space method. Zernike moments are used as a shape descriptor for complex shapes such as trademarks that are difficult to be defined with a single contour for similarity based image retrieval applications [1-2].

An important problem in pattern analysis is the automatic recognition of an object in a scene regardless of its positions, size and orientation. They arise in a variety of situations such as packaging of manufactured parts, classification of chromosomes, target identification, and scene analysis. The current approaches to invariant two-dimensional shape recognition include extraction of global image information using regular moments, boundary based analysis via Fourier descriptors, or auto regressive models, image representations by circular harmonic expansion, and synaptic approaches [3].
A fundamental element of all these schemes is definition of a set of features for representation and data reduction. Normally additional transformations are needed to achieve the desired invariant properties for the selected features. After invariant features are computed, they are compared with the offline database features for underlying images. However the aim of this work is to develop an efficient Content Based Image Retrieval System for the similar images which independent to rotation, translation and scaling.

2. EXISTING IMAGE RETRIEVAL SYSTEMS
The problem of retrieving and recognizing patterns in image has been investigated by many of the researchers for several applications. Some of them are as follows:

S. Alwis [3], Kato [4], and J. K. Wu [5] have proposed three of the most prominent trademark image retrieval systems. Different methodologies have been employed in these trademark retrieval systems. The Trademark system proposed by Kato uses graphical feature vectors (GF-vectors) to interpret the image content automatically and calculates the similarity based on the human perception [4]. The Star system adopts mainstream content based image retrieval technique (CBIR). The techniques involved in this method include Fourier descriptors, grey level projection and moment invariants. The Star system works by considering both shape components and the layout of the image [5]. However the recognition of some perceptually significant components has been considered to be too difficult to be done by automated process [4]. Therefore, to a certain extent, manual operation is needed for the segmentation of some abstract trademark images. The Artisan system, however introduces an innovative approach that incorporates principles derived from Gestalt psychology to cope with device only marks which consists of some abstract geometric design [6].

Apart from S. Alwis [3], Kato [4], and J. K. Wu [5] trademark image retrieval systems, a significant amount of research work has been also concentrated on trademark or logo image retrieval. The work of carried out by Graham [6] for some additional images of logo and trademarks. Hussain[7] have proposed the topological properties of self-organising map for the similarity retrieval from trademark image database. Cerri et al. [8] have considered geometrical-topological tools for describing trade mark shapes and matching their similarity based on size functions of the trademark images. Jiang et al. [9] have presented a new approach by using the adaptive selection of visual features with different kinds of visual salient features, including symmetry, continuity, proximity, parallelism and closure property of the trademark images. Hung et al. [10] exploited the contour and interior region for retrieving similar trademark images. Petrakis et al. [11] utilized relevance feedback for logo and trademark image retrieval on the web. Shen et al. [12] utilised block feature index to describe trademark shape. An enhanced normalization technique for the wavelet shape description of trademarks was developed by Li and Edward [13]. Kim and

Kim [14] have used Zernike moments as shape descriptors and conducted experiments based on the MPEG-7 core experiment procedure. Gaur et al. [15] proposed a mixed model based on three features for color, shape and DWT based CBIR system. He suggested that by considering these all features together ranking of the image database can be improved.

There are several remarkable image retrieval techniques that have been developed in recent years. This work is based on the Invariant image recognition by Zernike moments [1]. This particular approach has been used for black and white images and the problem for color images has been solved by using canny edge detector along with the Zernike moment algorithm.

3. OVERVIEW OF THE PROPOSED CBIR SYSTEM

The proposed content based retrieval system consists of an offline database construction part and an online retrieval part as shown in Fig. 1. The offline database construction is implied so as to ensure maximum retrieval efficiency by extracting a feature set i.e. Zernike moment and phase angle phi and storing this feature along with the corresponding image so that next time when a query image is presented before the system, the system does not need to go online for image retrieval.

To access the database the user initiates the online CBIR system by presenting a query image to the input, and then the system uses this image to extract the features from it. After that the system starts matching the features of the query image with the features stored in the database already. After the completion of feature matching the machine returns the results to the user.

3.1 Image pre-processing and edge detection

Sample images come in different sizes and shapes and their key components appear in different locations. A robust retrieval system should be invariant to the scaling, rotation and translation of sample images. To achieve such a robust system the proposed CBIR system performs the following steps to normalise each sample images:

- First the color image is changed to grayscale image for better edge detection
- Apply median filtering to remove salt and pepper noise, which is necessary only if noise is present.
- Then resizing is to be done before any further processing takes place.
- Apply canny edge detector to find the image discontinuities in an image: the canny operator produces binary edge map, which preserves the crucial structure properties and significantly reduces amount of data in an image. The resulting edge map represents the contour of the underlying sample image.

3.2 Feature Extraction

As shown in the figure 1 feature extraction is to be done both in offline mode as well as online mode for CBIR system. Feature extraction is about extracting distinctive features/or set of attributes that can describe the image in any condition. Weather the image is rotated by some degree, or it is scaled and resized, this particular feature should be invariant to all these mentioned factors which an image can face. The task of the feature representation module is to extract a set of features of the images from their corresponding edge maps generated by the Canny edge detector. The Zernike moments applied in the method of feature extraction to extract the global features of the images. Global feature extraction is presented in the successive sub-section.

3.2.1 Global Image Feature Extraction

Moments and function of moments can be utilized to identify pattern features in many applications [1]. These features utilize global information about the image and do not require boundary based descriptors as Fourier descriptors do. Regular moments are defined as [1]

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)dxdy \qquad (1)$$

Where m_{pq} is the $(p + q)^{th}$ order moment of the continuous image function f(x, y). In case of digital images the integral is replaced by summation and m_{pq} becomes

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y) \qquad (2)$$

The reason for selecting Zernike moments is that it is invariant to rotation i.e. rotating the image does not change the magnitudes of its Zernike moments. Thus they could be used as rotating image invariant feature for image retrieval process.

The Zernike moments are derived from a set of complex polynomials orthogonal over the interior of a unit circle U: x2 + y2 <= 1 and defined in the polar coordinates. The form of two-dimensional polynomial V_{mn} is expressed as

$$V_{mn}(\rho,\theta) = R_{mn}(\rho)\exp(jm,\theta) \qquad (3)$$

where n and m are called order and repetition, respectively.

The order n is a non-negative integer, and the repetition m is an integer satisfying n-|m| = an even number and |m| <= n. j is an imaginary unit √-1. $R_{mn}(\rho)$ is the one dimensional radial polynomial, which is defined as

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s (n-m)!\rho^{n-2s}}{s!(\frac{n+|m|}{2}-s)!(\frac{n-|m|}{2}-s)!} \qquad (4)$$

As the Zernike moments are the projection of image I(x,y) onto these orthogonal basis functions, the image can be decomposed into a weighted sum of the Zernike polynomials

$$I(x,y) = \sum_{n=1}^{\infty}\sum_{m=-n}^{n} A_{nm}V_{nm} \qquad (5)$$

where A_{mn} are the Zernike moments, i.e., the coefficients of the Zernike polynomials. The Zernike moments of an image I(x,y) with continuous intensity are calculated according to the following equation:

$$= A_{nm} = \frac{n+1}{\pi}\iint_U I(x,y)V_{nm}(\rho,\theta)dxdy \qquad (6)$$

For digital image of NxN pixels, the discrete form of the Zernike moments for an image expressed as:

$$= A_{nm} + \frac{n+1}{\lambda}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} I(x,y)V_{nm}(\rho,\theta) \qquad (7)$$

where $\lambda = \frac{\delta A}{\pi}$ is a normalizing constant,

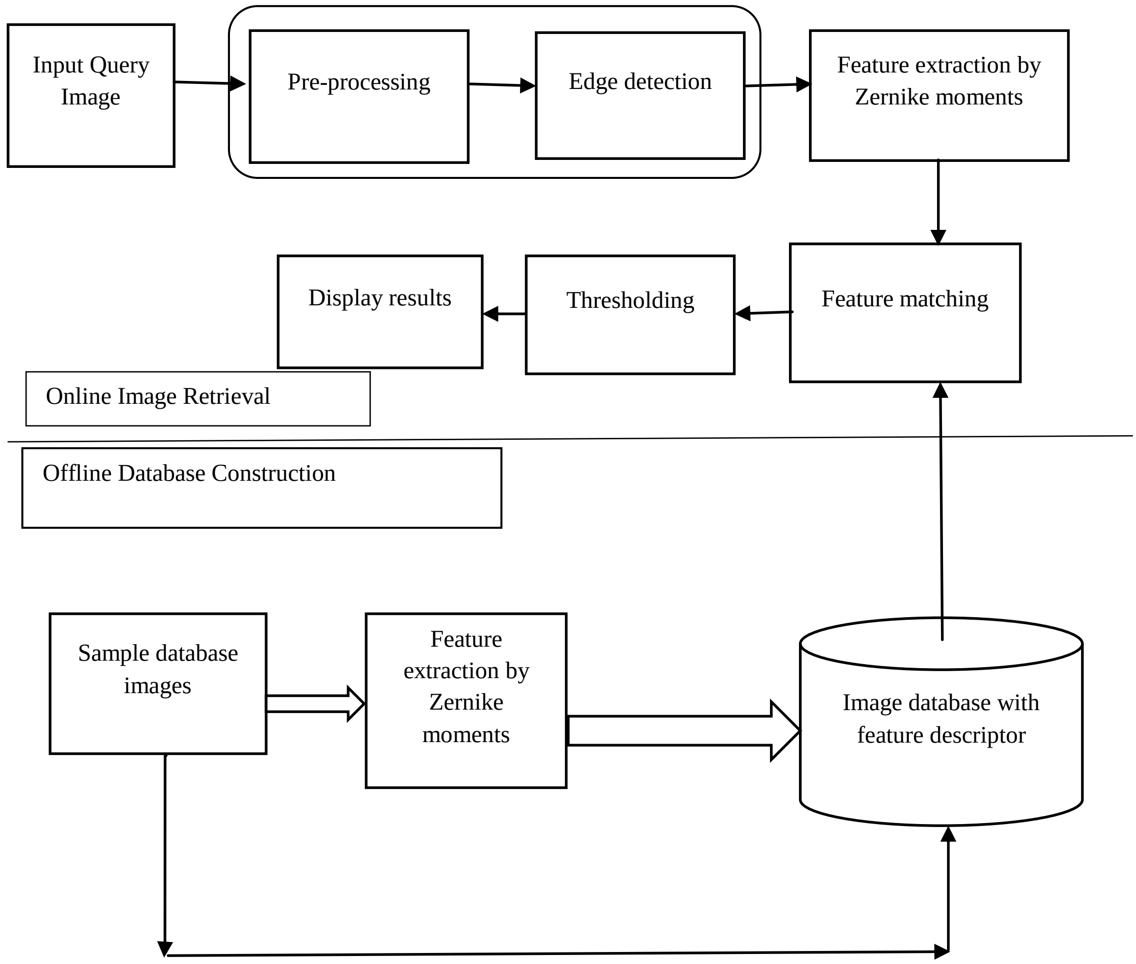


Fig: 1 The architecture of proposed CBIR system.

δA is the elemental area of the square image when projected onto the unit circle of the Zernike polynomials. As the Zernike basis functions take the unit disk as their domains, the disk must be specified before the moments are calculated. In this work, all the sample images are extracted and projected onto a unit circle of fixed radius of 200 pixels. This particular step makes the resulting moments invariant to translation and scaling, in addition to rotation-invariant nature possessed Zernike polynomials. The purpose of using Zernike moments is to explore an image into a series of orthogonal bases; the precision of representing the accurate shape of the image depends upon the number of moments used from the expansion. For the case of human eye higher orders i.e. orders greater than five or six are too small to be measurable reliably so higher orders are always ignored preferably [14]. In this work the order of the moments are restricted to 4-order Zernike moments so as to yield the most efficient and effective measurement of the global shape.

3.3 Component Feature Matching

Feature matching is about measuring the similarity between feature vectors of the query image and the database images. An appropriate choice of feature matching technique can enhance the system performance while an inappropriate matching approach can lead to unexpected results from the system even though an effective and robust system has been deployed for feature representation.

The sample images can be structurally and conceptually alike with different interior details, during the feature matching stage. Therefore the global features are compared separately in order to enhance the system performance. The reader must be reminded that the global feature components include 15 Zernike moments of the orders 0-4.

By utilizing the Euclidean distance to compute the similarity between the query image features and the features stored in the database, the similarity between two images can be obtained. The threshold for the Euclidean distance is kept 0.006 for 20% matching and is normalized in the range [0, 1]. The purpose of keeping this method for similarity matching is, if the query image Zernike moment and the database image Zernike moment’s Euclidean distance is below the threshold then the image is kept under the category of relevant image. And if the Euclidean distance is more than the prescribed threshold, the query image comes under the category of irrelevant image and a message is displayed “image not in the database”. If the distance between the query image and database image is greater than 0.006 for the moment feature, a penalty of 1 is added to Euclidean distance value. The final value obtained for the Euclidean distance for each image can be used to rank images in ascending order. The query images with the minimum distance are placed on the top while the images having larger Euclidean distance are placed at the bottom. Then with the selection of proper threshold for similarity criteria only similar images can be displayed.

4. RESULTS ANALYSIS

In this section, the results obtained using the proposed algorithm under various conditions of orientation are presented in table 1 and table 2. Zernike moments were calculated and compared using Euclidian distance for performance evaluation. Through the tabular results it is clear that:

➢ For the first part of result a color image, which consists of text and picture have been taken to show performance of the proposed technique. Results in the first and second table indicate that Zernike moments goes on changing if the size of image reduces to 25% of its original size. Otherwise it remains same.

➢ The results generated by the color and texture features [15] method have been shown in the figure 4. It is clear from the images that the images with closer values of

➢ distance are sorted closer to each other rank wise and the images with greater distance have been sorted out and ranked further away.

In this paper Lena image was taken as a reference for identification and the main purpose was to identify the similar images of Lena from the database having different types of images including images of a different girl with different textures and images of vegetable and animals. However the color, shape and DWT [15] was able to classify and identify the images similar to Lena. But some of the images that were classified and sorted out were not the Lena images, instead they were the images of some different girl. Whereas the proposed approach solves the issue of classification and retrieval of the exact Lena images fig. 2 with different orientations, and retrieved only image of Lena in the dataset. The values of the Zernike moments have been mentioned along with the respective images in fig. 5.

Fig. 2 Query Image_1.

Fig. 3 Query Image_2

Fig. 4 Results for the query Lena image considering color and texture features

Fig. 5 Considering only Zernike moment.

Table 1. Zernike moments with different orientations for 200x200 image size

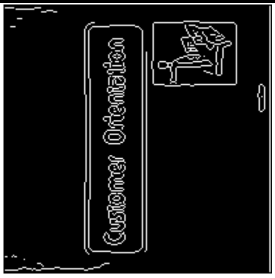
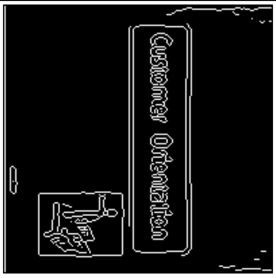
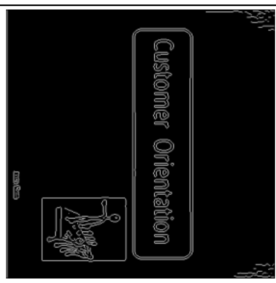
Angle of Rotation	Zernike Moments for image size 200X200	Angle of Rotation	Zernike Moments for image size 200X200
0	 A= 0.051328 ϕ = -75.3398	90	 A= 0.051005 ϕ = 4.9082
180	 A = 0.050951 ϕ = -75.1264	270	 A = 0.051158 ϕ = 4.741

Table 2. Zernike moments with different orientations for 400x400 image size

Angle of Rotation	Zernike Moments for image size 200X200	Angle of Rotation	Zernike Moments for image size 200X200
0	 A = 0.27969 ϕ= -177.3398	90	 A = 0.27969 ϕ= -2.2898
180	 A = 0.27969 ϕ= -177.7667	270	 A = 0.27969 ϕ= -2.2898

5. CONCLUSION

The performance of the Zernike algorithms as a reliable shape descriptor is compared with a better combination of Zernike algorithm with canny edge detector for color images. It is seen that this combination of edge detector is producing better results for color images. It was also seen that if the Zernike algorithm was applied without any edge detector for color images the value of moments were not invariant to some extent with respect to the orientation of the image. Instead a Zernike algorithm with an edge detector is applied and the values of the Zernike moments for various orientations for a 200X200 images and 400X400 images are evaluated, giving a better and stable values for Zernike moments.

6. REFERENCES

[1] Y. S. Abu-Mostafa and D. Psaltis, "Image Normalization by Complex Moments," IEEE.

[2] S. A. Dudani, K. J. Breeding, and R. B. McGhee, "Aircraft Identification by Moment Invariants," IEEE Trans. Computer, vol. C-26, no.1, pg. 39-45, Jan. 1983.

[3] S. Alwis, J. Austin, "A Novel Architecture for Trademark Image Retrieval Systems," Proceedings of the Challenge of Image Retrieval, British Computer Society, UK, 1998.

[4] T. Kato, "Database Architecture for Content-Based Image Retrieval," Proc. SPIE Image Storage Retr. Syst. 1662, pg. 112-123, 1992.

[5] J. K Wu, C. P. Lam, B. M. Mehtre, Y. J. Gao, A. Narasimhalu, "Content-based Retrieval for Trademark Registration," Multimedia Tools Appl. 3 (3), pg. 245-267, 1996.

[6] J. P. Eakins, M.E. Graham, J. M. Boardman, "Evaluation of a Trademark Image Retrieval System", Information Retrieval Research, the, 19th Annual BSC-IRSG Csolloquium on IR Research, 1997.

[7] M. Hussain, J. P. Eakins, Component-based Visual Clustering using the Self-Organizing Map, Neural Networks 20 (2), pg. 260-273, 2007.

[8] A. Cerri M. Ferri, D. Giorgi, "Retrieval of Trademark Images by Means of Size Functions", Graphical Models 68 (5-6) pg. 451-471, 2006.

[9] H. Jiang, C. –W. Ngo, H. –K. Tan, "Gestalt-based Feature Similarity Measure in Trademark Database," Pattern Recognition 39 (5), pg. 988-1001, 2006.

[10] M. H. Hung, C. H. Hsieh, C. M. Kuo, Similarity Retrieval of Shape Images based on Database Classification," J. Visual Commun. Image Representation 17 (5), pg. 970-985, 2006.

[11] E. G. M. Petrakis, K. Kontis, E. Voutsakis, "Relevance Feedback Methods for logo and Trademark Image Retrieval on the Web", Proceedings of the 2006 ACM Symposium on Applied Computing, pg. 1084-1088, 2006.

[12] D. F. Shen, J. Li, H. T. Chang, H. H. P. Wu, "Trademark Retrieval based on Block Feature Index Code," Proceedings of IEEE International Conference on Image Processing 2005, vol. III, 2005, pg. 177-180.

[13] Q. Li, J. Edwards, "An Enhanced Normalization Technique for Wavelet Shape Descriptors", The Fourth International Conference on Computer and Information Technology, 2004, pg. 722-729.

[14] W. Y. Kim, Y. S. Kim, A Region-based Shape Descriptor using Zernike moments, Signal Process. Image Commun. 16 (1-2), pg.95-102, 2000.

[15] Sanjay Gaur and Rashmi Sabu. "The Performance Evaluation for CBIR Using Colour, Texture and DWT Feature for Image Retrieval", international Journal of Bussiness & Engineering Research, Volume 7, pg. 1-9, 2013.