International Journal of Electrical and Computer Engineering (IJECE)
Vol. 7, No. 3, June 2017, pp. 1651~1660
ISSN: 2088-8708, DOI: 10.11591/ijece.v7i3.pp1651-1660

Shape based Image Retrieval using Lower Order Zernike Moments

G. Sucharitha, Ranjan K. Senapati
Department of Electronics and Communication, K L University, India

Article Info

Article history:

Received Aug 9, 2016
Revised Oct 28, 2016
Accepted Nov 12, 2016

Keyword:

CBIR
Euclidian distance
GFD
LZM
ZM

ABSTRACT

Shape is one of the significant features of Content Based Image Retrieval (CBIR). This paper proposes a strong and successful shape feature, which is based on a set of orthogonal complex moments of images known as Zernike moments. For shape classification Zernike moment (ZM) is the dominant solution. The radial polynomial of Zernike moment produces the number of concentric circles based on the order. As the order increases number of circles will increases, due to this the local information of an image will be ignored. In this paper, we introduced a novel method for radial polynomial where local information of an image given importance. We succeeded to extract the local features and shape features at very a low order of polynomial compared to the state of traditional ZM.The proposed method gives an advantage of a lower order, less complex, and lower dimension feature vector.For more similar images we find that simple Euclidian distance approximately zero. Proposed method tested on a MPEG-7 CE-1 shape database, Coil-100 databases. Experiments demonstrated that it is outperforming in identifying the shape of an object in the image and reduced the retrieving time and complexity of calculations.

Corresponding Author:

G.Sucharitha,
Departement of Electronics and Communication Engineering,
KL University,
Green Fields, Vaddeswaram, Guntur (Dist) 522502, Andra Pradesh, India.
Email: sucharithasu@gmail.com

1. INTRODUCTION

The process of image retrieval depends on various image features. The features represent the image more significant. Features include color, texture, shape, shadows, etc. In all features shape of an object gives a powerful clue in identifying the object [1]. So, an effective and efficient shape descriptor is required. The shape descriptors are divided into two categories, contour based and region based [2]. Contour based shape descriptor gives the boundary information of an object. Here it ignores the interior information of the shape. The algorithms used for this are Fourier descriptors, Elongation, Area, Curvature scale space etc [3]. The region based descriptors make use of boundaries and interior regions of the shape. In region based methods, the information captured from both boundaries and interior region of shape. Moments are the commonly used approaches for shape identification.

In general, the moments are quantitative values that describe a distribution, raising the components to different powers [4].The definition of the regular moment or geometric moment is

$$M_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\,dx\,dy$$

Where $f(x,y)$ is a real function, p & q are order of a polynomial.

However, geometric moments contain highly redundant information. As the order increases computational complexity increases. It doesn't have any desired rotational invariance, translation and scale invariant. In the geometric moments the basis function $x^p y^q$ is not orthogonal. The reconstruction of image with these moments would involve redundant information among all the moments and produce high computational complexity.

A. Khotanzad and M.R. Teague et.al. [5], [6], introduced a two dimensional orthogonal moments i.e Zernike moments, which are invariant to image translation, orientation and size. It has been observed that the magnitude of ZM would not change for any rotation and scaling of an image. Due to these properties of ZM, it outperforms than many other descriptors, such as Geometric, Legendre and Pseudo Zernike moments [7], [8]. Nugroho and Tian et.al [9], [10] proved that, the rotarional invariant property gives the dominant results compare to other shape descriptors.

However, direct computation of these moments in higher order is very expensive. There is a great need to limit their use at higher orders. The cause behind is not only a computational complexity, but also highly sensitive to noise [11], [12]. The performance of the system may diminish if the order and moment are chosen properly.

In this paper we have not only tried to reduce the order, but also tried to reduce the dimensions of the feature vector. Generally the order of the Zernike polynomial decides the No.of concentric circles. As the polynomial order increases, the No.of circles also increases. So, the area covered by two consecutive circles will be minimized. If this area is minimized, the shape with interior information of an object is not possible, and with more number of moments there is a high similar information which is not much useful while retrieving.In the proposed method the aspects are taken into consideration: order and dimensions of feature vector.

The remainder of the paper is organized as follows. Section 2 introduces ZMs, Section 3 describes Image mapping and transform diatance, the proposed method is in Section 4, experimental results and comparison are presented in Section 5. Finally, in Section 6conclusion are outlined.

2. ZERNIKE MOMENTS

Complex Zernike moments constitute a set of orthogonal basis functions mapped over a unit circle. The orthogonal property of ZM's suits better for shape recognition schemes. This property shows that the contribution of each moment is independent and unique. Due to this property the redundancy has been reduced as minimum as compared to the geometric moments.

Mathematically, Zernike basis function is defined with an order *n* and repetition *m* over C={(n,m)| 0≤ n≤ ∞, |m|≤ n, |n-m|=even}.

$$Z_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho,\theta) V_{nm}(\rho,\theta)\, d\rho d\theta \quad (1)$$

where

$$V_{nm}(\rho,\theta) = R_{nm}(\rho).* e^{-jm\theta} \quad (2)$$

and

$$R_{nm}(\rho) = \sum_{k=0}^{(n-|m|)/2} (-1)^k \frac{(n-k)!}{k!\left(\left(\frac{n+|m|}{2}\right)-k\right)!\left(\left(\frac{n-|m|}{2}\right)-k\right)!} \rho^{n-2k} \quad (3)$$

Where *n* is a positive integer representing the order of the radial polynomial and *m* is no.of repetitions. Where *f (x, y)* is a function of an image with the size of NxN. For digital images the integrals in Equation 1 are replaced by summations.

$$Z_{nm} = \frac{n+1}{\pi} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y).* V_{nm}(\rho,\theta) \quad (4)$$

3. IMAGE MAPPING AND TRANSFORM DISTANCE

The orthogonality and completeness of Zernike polynomials allow us to represent any square image function into the unit disk [8] like as shown in Figure 1.

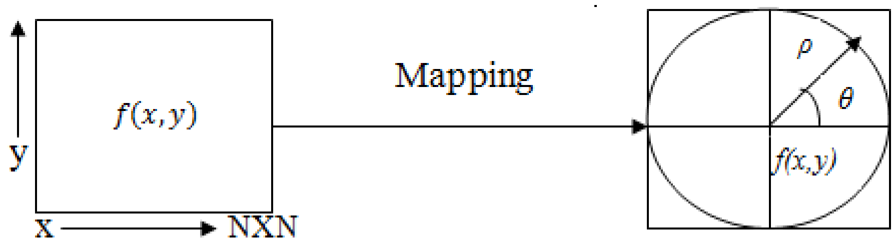


Figure 1. Transforming the image into unit circle

The transformed distance ρ and the phase angle θ at a pixel (x,y) are designed in such way to insert the image into the unit disk. The equations are

$$\rho = \frac{\sqrt{(2x-N+1)^2+(2y-N+1)^2}}{N} \tag{5}$$

$$\theta = tan^{-1}(\frac{2y-N-1}{N+1-2x}) \tag{6}$$

To map the digital image into the circle, first the image has to convert into a square image i.e NXN, where N is even.

As a matter of fact, if the image size is odd (i.e N), the center of the image is $(x, y) = \left(\frac{N+1}{2}, \frac{N+1}{2}\right)$ at this point from the above equations $\rho = 0$ and the phase angle $\theta = Tan^{-1}\left(\frac{0}{0}\right) = Nan.$

The way to resolve this problem is to select the image size is square as well as even matrix. The image not have any center and no redundancy.This transform distance generate some concentric circles for any kind of NXN matrix. Here we are bringing the ring areas, which denotes the image transformed into the circle in polar coordinate system as shown in Figure 2. Let ***o*** be the center of an object and ***R*** be the incremental radius length. Using the Equations (5), (6) rings were constructed from the center to the boundary of an object. The area within the first ring constitutes the first ring area and the area between first and second is a second ring area. The human observation results, the pixel values in a particular ring area are approximately equal, but the pixel values linearly increase from the center to the outward ring. The outermost ring pixel values are greater than or equal to unity. The innermost ring, i.e at the center pixel value is completely zero. With this $\boldsymbol{\rho}$**,** the lower order Zernike moments are incapable to give the shape of an object.

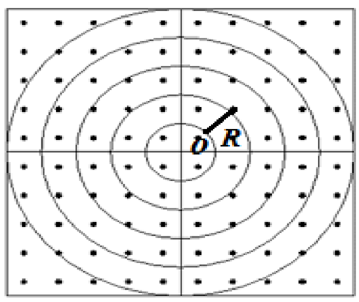


Figure 2. Ring areas

To overcome this problem, the transform distance for a NXN matrix is designed in such way that, the pixel values are customized in their respective rings. Assume that there are C number of rings and R denotes the distance between two adjacent rings.

The transform distance $\boldsymbol{\rho}$ matrix

$$\rho(x) = \begin{cases} 1 & x \leq R \\ x + \alpha(r_{C_1}) & R < x \leq 2R \\ . & . \\ . & . \\ x + \alpha(r_{C_c}) & (C-1)R < x \leq CR \end{cases} \quad \textbf{(7)}$$

The normalized function $\boldsymbol{\alpha(r)}$, for a ring area can be defined as

$$\alpha(x) = \begin{cases} 1 & x \leq R \\ \frac{C-2}{C} & R < x \leq 2R \\ . & . \\ \frac{C-k}{C}(k-1) & R < x \leq kR \\ . & . \\ . & . \\ 0 & (C-1)\,R < x \leq CR \end{cases} \quad (8)$$

For the orders shown in Table 1, the radial polynomial calculated using above transform distance method. The practical observations cleared that, the number of circles for any NXN matrix are 10 for the polynomial order lessthan or equal to 5. From the above calculations the innermost and outermost circle pixel values are 1and 0 respectively, so the pixel values above the outermost circle are zero. Using the Equations (2), (3) and (4) the radial polynomial, and Zernike moments were calculated for the orders shown in Table1.

4. PROPOSED LZW

The proposesd block diagram shown in Figure 3 with various steps involved in calculating the feature vector. The entire database as well as the query image has to resize in order to get a square size matrix. Since, to calculate Zernike moments the image should in square size. Here we followed a practice of resizing (i.e interpolation) based on the dimensions of images.

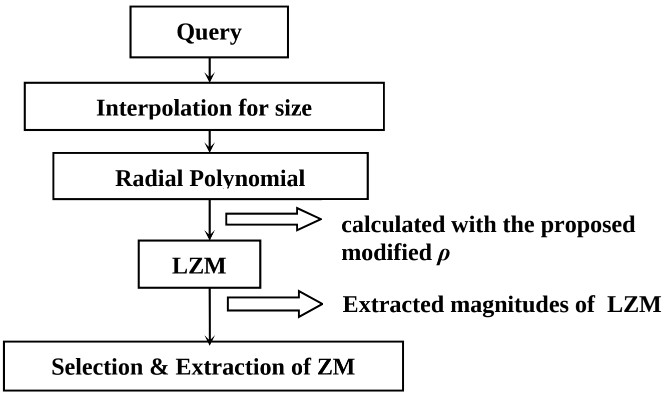


Figure 3. Block diagram of proposed method

For query image and the database images LZMs are calculated. The transform distance, the radial polynomial $R_{nm}(\rho)$ plays a main role in extracting the shape features. The order of the polynomial decides the number of non zero circles and its area. As this order goes on increasing, the number of concentric circles also increases. As a commom thought that, if the number of circles increases of a given space, the area covered by each ring (i.e area between two concentric circles) will reduce.

In order to analyze the effect of orders, the moments are divided into two groups. The two groups are

Group 1: $$Z_{nm} = \begin{cases} 3 \leq n \leq 5 \\ |m| \leq n \end{cases} \quad (9)$$

Group2: $Z_{nm} = \begin{cases} 6 \leq n \leq 15 \\ |m| \leq n \end{cases}$ (10)

The feature vector consists of only magnitudes of Zernike moments. The simple metrics used to find the similar images from the database are Euclidian, Manhattan distance. If all the moments Z_{nm} of the image *f(x,y)* with an order *n* are known, it is proved that reconstruction of an image is possible. The reconstructed function can be formed as follows:

$$f(x,y) = \sum_{n=0}^{n_{max}} \sum_{m=-n}^{n} Z_{nm} V_{nm}(\rho,\theta) \quad (11)$$

Table1. List of Lower Order Zernike Moments for Group1

Order	Zernike moments	Dimensionality of the specified order
0	$Z_{0,0}$	1
1	$Z_{1,1}$, $Z_{1,-1}$	2
2	$Z_{2,0}$, $Z_{2,2}$, $Z_{2,-2}$	3
3	$Z_{3,1}$, $Z_{3,-1}$, $Z_{3,3}$, $Z_{3,-3}$	4
4	$Z_{4,0}$, $Z_{4,2}$, $Z_{4,-2}$, $Z_{4,4}$, $Z_{4,-4}$	5
5	$Z_{5,1}$, $Z_{5,-1}$, $Z_{5,3}$, $Z_{5,-3}$, $Z_{5,5}$,$Z_{5,-5}$	6

4.1. Feature Extraction

From the preceding, we obtained scale invariant, rotational invariant and translational invariant ZM features. However, from the table1 the number of moments calculated for group1 is 21. But all these features are not required for shape identification. The values of Z_{00} and Z_{11} are constant for all normalized and binary images, they are not included as image features. The experiments' results show that ZMs, with the max order up to five, could have a sufficiently good image representation power.

4.2. Query Matching

Feature vector for query image is represented as $I_q = (I_{q1}, I_{q2}, I_{q3}, \ldots . I_{qn})$ is obtained after the feature extraction. Similarly, each image in the database is represented with their own feature vectors $I_{DB_i} = \{I_{DB_{i1}}, I_{DB_{i2}}, I_{DB_{i3}}, \ldots . I_{DB_n}\}$, $i = 1,2,3 \ldots |DB|$. The aim is to retrieve the possible best images that resemble the query image. This involves selection of some top matched images by measuring the distance between query and database images.

Euclidian distance is used as a similarity metric. To find the similarity only the magnitude of LZMs are considered. The distance metric results will be used in retrieving the images from the database by sorting and ranking. The Euclidian distance measured using following formula.

$$Distance = \frac{\left(\sum_{j=1}^{n} \left(I_{DB_{ij}} - I_{qj}\right)^2\right)^{1/2}}{N} \quad (7)$$

Where *n*is the length of feature vector, N is the number of images in the database.

4.3. Experiments

To evaluate the overall performance of the proposed image retrieval method based on shape and local features , the MPEG-7 CE-1, Coil-100 databases are used. Some examples form MPEG-7CE-1 database as shown in Figure 4.

Experiment *1

In this experiment, MPEG-7 CE-1 database [13] used. This database consists of a large number of binary images of various shapes of different objects in various orientations. These binary images are pre-classified into different categories, each of size is 20 and total number of objects are 70 by domain professionals.

Rotation test: There are 70 groups of images where each group has 20 similar images with different orientations. All the 70 images from all groups are used as queries.

***Noise test**:* The efficiency of the descriptor verified under rotation, scaling, reconstruction with different noise added images. Various levels of salt & pepper, Gaussian noise is used with various SNR values as shown in Figure 7.

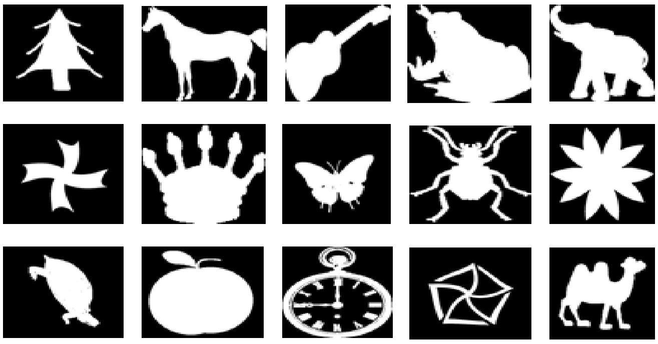


Figure 4. Examples from MPEG-7 CE-1 Database

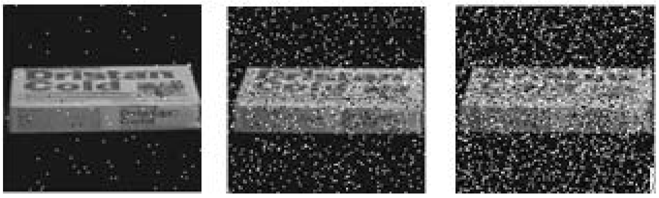


Figure 5. Sample images with salt& pepper noise under different values of SNR

Subject test: This database has 1400 binary images, with 70 subjects and each subject has 20 similar images with all rotations and scaling images. Fig6 shows some samples of the MPEG-7 database. All 70 subject images used as queries.

Experiment *2

In this experiment Coil-100image database is used [14]. Coil-100 database having 100 different objects with each object has 72 similar images with different orientations. This database helps to verify the rotational property of the descriptor in all angles.All 100 images from various objects used as queries.Some samples from Coil-100 database shown in Figure 6 with various rotations. All the above tests performed on these databases also.

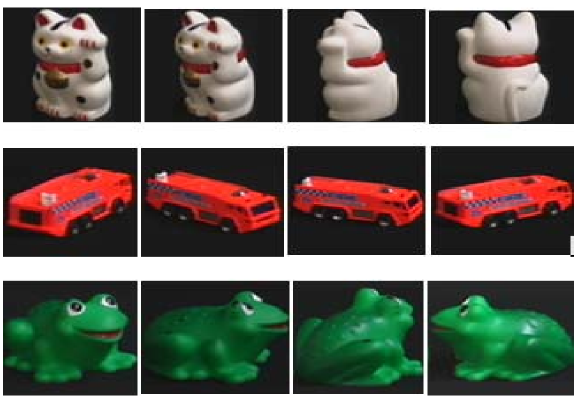


Figure 6. Sample images from Coil-100 Database

To evaluate the proposed descriptor's performance in real images, Coil-100 database has been chosen. To prepare for our experiments, the images converted into gray scale images and selected six views of different orientations per object. The selected object from the database can be considered as a combination of scaling, rotation and subject test database. Figure 9 and Table 2 shows the comparison and precision, recall calculations.

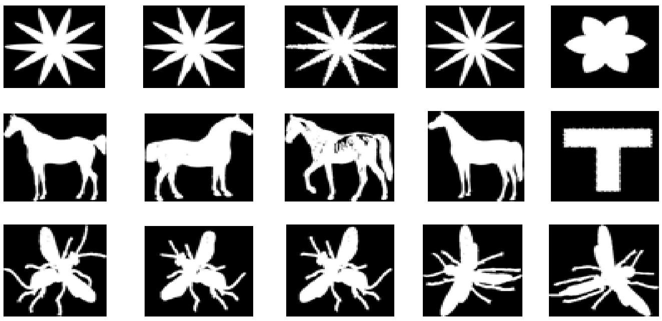
In all experiments, each image in the database is used as the query image. For each query, the algorithm collects a set of similar images from the database $I_x = \{I_{1,}I_2,, \ldots .I_z\}$ with the shortest image matching distance computed using Equation (7). The performance of the proposed method is measured in terms of precision and recall as shown below.
The precision and recall defined as

$$Precision(P) = \frac{NumberofrelevantImagesretrieved}{TotalNumberofimagesretrieved} \quad (8)$$

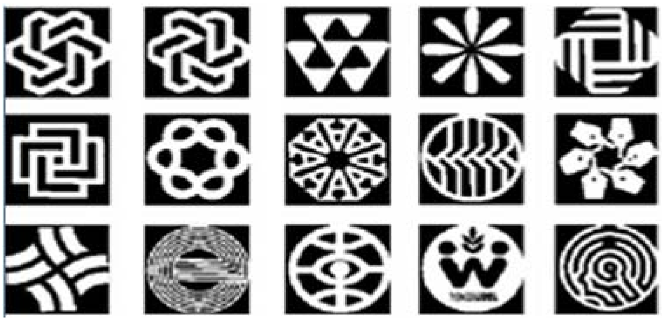
$$Recall(R) = \frac{Numberofrelevantimagesretrieved}{TotalNumberofrelevantimagesinthedatabase} \quad (9)$$

5. RESULTS

A number of experiments were performed to appraise the performance of the proposed ZM, named LZM (Lower order Zernike moment) with corrected weights in the transform distance circles. These moments compared to commonly used ZMD with the order of 10, and Generic Fourier Descriptor (GFD).



(a)



(b)

Figure 7. Retrieval performance of LZM and ZMD_{10}, (a) Retrieved results using LZM, (b) Retrieved results using ZMD_{10}. The query image corresponds to top left of each row.

Figure 7 shows the retrieval results for the proposed method and ZMD in the order of ten. For instance, LZM returned with similar images with different rotating images also. Fo Figure 8 shows the retrieved results for LZM and GFD on Coil-100 database. The proposed method returned top five similar images in different orientations. GFD gave a result with one or two unmatched images in the top five list of images. Figure 9 is giving a precision, recall curves for proposed method (LZM) and ZMD_{10}, GFD on Coil-

100 database. Figure 10 is showing P-R performance of a proposed method for various SNR values. The graph is based on the results of a noise image query and its respective retrieved images from the original database. From the Table 2: we can observe the percentage of precision and recall values for number of images per category are 20 and 30. The shown P-R values are for single threshold for all categories (approximately 0.09). From the results it is concluded that, the proposed method identifying the shape, including local information of an image.

(a)

(b)

Figure 8. Top five retrieved images using LZM and GFD on Coil-100 database. First image of the each row is query image.

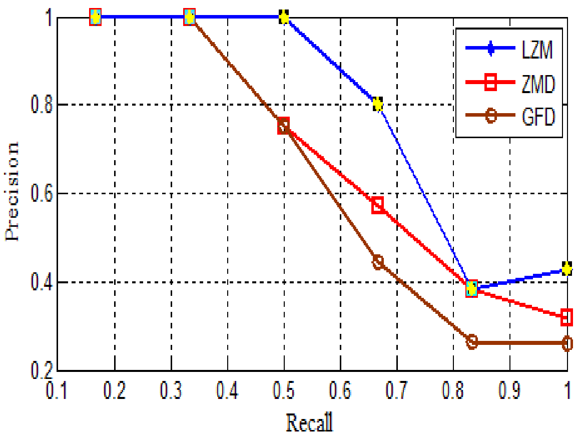


Figure 9. P-R performance results based on using, separately, LZM, ZMD,GFD on the Coil-100 database

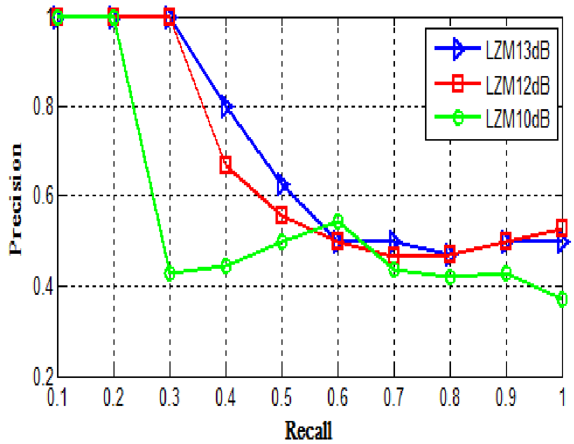


Figure 10. P-R performance of proposed method for various SNR values

Table 2. For the first 10 categories of Coli-100 database precision and recall values for first n=20 and n=30 images per category

Category	n=20 Precision (%)	n=20 Recall (%)	n=30 Precision (%)	n=30 Recall (%)
1	78.9	75	62.06	60
2	90.9	100	90.6	96.6
3	16.9	45	14.5	33.3
4	25.9	100	27.5	100
5	22.2	100	24	100
6	38.2	65	32.14	60
7	21.05	100	21.8	100
8	40	70	42.3	73.3
9	100	95	100	70
10	71.4	50	91.6	36.6

6. CONCLUSION

In this paper, we introduced an efficient shape feature descriptor with less complexity and redundancy. The transform distance matrix designed in such a way, in assigning the proper weights to the respective circle pixel values. The base function $R_{nm}(\rho)$ calculted using this proposed ρ function. Through the LZM we capture the maximum information about the image for a small no.of polynomial order. Since the LZW feature vector length aproximately 18, for n=5. The results proved that, praposed methos is scale ivariant and rotational invariant. The proposed method proved its efficiency in retrieving the images from the database, even for a noisy query image.

REFERENCES

[1] Liu, M., He, Y., & Ye, B, (2007), "Image Zernike Moments Shape Feature Evaluation based on Image Reconstruction", *Geo-spatial Information Science*, *10* (3), 191-195.

[2] Bober, M. (2001), "MPEG-7 Visual Shape Descriptors", *IEEE Transactions on Circuits and Systems for Video Technology*, *11*(6), 716-719.

[3] Mehtre, B. M., Kankanhalli, M. S., Lee, W. F. (1997), "Shape Measures for Content based Image Retrieval: A Comparison", *Information Processing & Management*, 33(3), 319-337.

[4] Vorobyov, M, (2011), "Shape Classification using Zernike Moments", Technical Report, iCamp-University of California Irvine.

[5] Khotanzad, A., Hong, Y. H., (1990), "Invariant Image Recognition by Zernike Moments", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 12(5), 489-497.

[6] Teague, M. R., (1980), "Image Analysis via the General Theory of Moments", *JOSA*, 70(8), 920-930.

[7] Hitam, M. S., Bakar, S. A., Jawahir, W. N., Yussof, W., "Content-Based Image Retrieval using Zernike Moments for Binary and Grayscale Images".

[8] Li, S., Lee, M. C., Pun, C. M., (2009), "Complex Zernike Moments Features for Shape-based Image Retrieval", *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 39(1), 227-237.

[9] Nugroho, Saptadi, Darmawan Utomo, "Rotation Invariant Indexing for Image using Zernike Moments and RTree", *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 9.2 (2013): 335-340.

[10] Tian, Gao, Wang Qi, "Target Recognition Algorithm based on BP Networks and Invariant Moments", *Indonesian Journal of Electrical Engineering and Computer Science*, 11.2 (2013): 969-974.

[11] Tahmasbi, A., Saki, F., Shokouhi, S. B. (2011), "Classification of Benign and Malignant Masses based on Zernike Moments", *Computers in Biology and Medicine*, 41(8), 726-735.

[12] Amayeh, G., Erol, A., Bebis, G., Nicolescu, M. (2005), "Accurate and Efficient Computation of high Oder Zernike Moments", In *Advances in Visual Computing* (pp. 462-469).

[13] Springer Berlin Heidelberg, http://www.cis.temple.edu/latecki/TestData/mpeg7shapeB.tar.gz

[14] S. A. Nene, S. K. Nayar, H. Murase, "Columbia Object Image Library (COIL-100)," Dept. Comput. Sci., Columbia Univ., New York, Tech. Rep. CUCS-006-96, 1996.

BIOGRAPHIES OF AUTHORS

G. Sucharitha received B.Tech degree in Electronics and Communication Engineering from JNTU, India, 2005 and M.Tech degree in Digital system and computer electronics from JNTU, India, 2009. Presently pursuing Ph.D from KL University, India. She has 11 years of teaching experience and presently working as an Asst.Professor in FST, IFHE University, Hyderabad, India. Her interested areas are Image Processing, Signal Processing and Digital Electronics.

Ranjan Kumar Senapati is a professor in KL University, India. He received his M.Tech degree forn NIT, Warangal, India, 2004. He received his Ph.D degree from Electronics and Comm. Engineering, NIT Rourkela, India, 2013. He has more than 12 years of teaching experience and his interested areas are Image Processing: Image Compression, VLSI implementation of image compression algorithms, Image watermarking, Image Analysis.