Applied Mathematical Sciences, Vol. 9, 2015, no. 47, 2325 - 2335
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2015.53214

Affine Invariant Compact Centroid Distance Shape Descriptor for Image Retrieval

P. Arjun

University College of Engineering Villupuram
PO Box 605 103, Tamilnadu, India

T. T. Mirnalinee

Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam
PO Box 603 110, Tamilnadu, India

Abstract

Simple and fast feature extraction methods are in need today for Content Based Image Retrieval (CBIR) and object recognition applications. The work presented in this paper is contour based one dimensional shape feature extraction technique for closed contour objects. The continuous contour is normalized into 'N' representative points. The sector area based object area normalization (OAN) technique is used for contour normalization. The centroid distance from all normalized points forms 1-D Compact Centroid Distance (CCD) feature vector. The experiments are conducted in MPEG-7 CE Shape-1 Part-B dataset images to test affine invariance property and image retrieval accuracy. Experimental results show the effectiveness of the proposed method in content-based image retrieval tasks.

Keywords: Centroid distance, object recognition, objects area normalization, shape feature extraction

1 Introduction

The image analysis tasks get more significance due to the exponential growth of image databases used in various domains like advertising, engineering design, machine vision, medicine, education, web searching, defence, remote sensing etc.

In image analysis, the useful information present in the image is extracted which describes the visual contents of the image in the form of numerical feature vectors. The primitive features extracted from the images are colour, texture and shape. These low level visual features are widely used in CBIR systems and object recognition applications [12]. The state of art image analysis systems are lacking in contrast with human visual system because the low level features in machine vision applications are unable to find the semantic information in the images. This 'semantic gap' motivates the researchers to develop advanced and efficient approaches to make the image users more comfortable.

The querying in image search engines exists today are heavily depends upon text based keywords. The keyword based image retrieval [9] is based on annotated keywords of the image and the metadata associated with the image and it ignores the visual contents present in the image. The limitations in this type of image retrieval are it needs more human annotation, wrong annotation causes irrelevant results and some image contents are very difficult to describe by keywords [6]. These limitations are motivating automatic retrieval (CBIR) of images using image features without any human intervention.

The object recognition task is also similar to CBIR in the sense that it extracts image features and identifies a given object in the image. A good object recognition algorithm [3] works analogous like to human visual ability to recognize objects in the scene that are in different viewpoints and in different lighting conditions. The object shape deformations like viewpoint changes and movement of object in real world situation are approximated in computer vision experiments by using geometric invariant transformations called affine transformations [5], [13].

In this work, a contour based affine invariant shape feature extraction method is proposed to describe objects present in the image. The object area normalization method [5] normalizes the contour into 'N' representative points using sector area approach. The CCD feature vector formed is the distance between the centroid to 'N' normalized points. This shape descriptor is simple, fast and more accurate than basic centroid distance descriptor. The MPEG-7 CE Shape-1 Part-B dataset images [10] are utilized for experimental evaluation. The correlation coefficient metric is applied on feature vectors of query, affine transformed and dataset images to retrieve the relevant images from the dataset.

2 Related works

The feature extraction plays vital role in all image analysis applications. Among the three primitive features the shape is the foremost feature to describe the objects in an image and more than 71% of image users are interested in using shape features for image retrieval tasks [2]. The major classifications of shape feature extraction methods are region-based and contour-based [7]. The region based approaches accounts all the pixels of the shape but in contour based approaches the boundary pixels of the shape alone is considered. The shape descriptor discussed in this paper is contour-based.

The simple and fast one dimensional descriptors for binary shape objects like area, centroid, axis of least inertia, eccentricity, perimeter, thinness ratio, irregularity, aspect ratio, Euler number, etc are used in QBIC system developed by Flickner et al. [6]. The 1D shape signatures like curvatures, tangent angles, complex co-ordinates, centroid distance etc are used in Fourier Descriptor (FD) based image retrieval systems [11].

The default shape descriptor for for MPEG-7 standard is Curvature Scale Space (CSS) descriptor developed by Abbasi et al. [1]. In CSS the contour is first normalized using equal arc-length parameter then Gaussian filter smoothes the contour of varying scales which produces CSS image. The maxima in CSS image describe the convexity/concavity of the shape.

The Beam Angle Statistics (BAS) descriptor [4] normalizes the continuous contour into 'N' normalized points using K-Curvature function. From all normalized points on the contour, an angle value is computed between two adjacent equal distance points (pair of bearings). Similarity between the shapes are computed using elastic matching algorithm. This stochastic shape descriptor is Rotation, Scaling, and Translation (RST) invariant.

The object area normalization based shape descriptors partition the total object area into equal part area segments using triangle area approach [13] or sector area approach [5] with respect to centroid. They are using Euclidean distance or correlation coefficient for matching the shapes. These shape descriptors satisfy affine invariance property and robust to shape deformations and noise.

Alajlan et al. [2] develops a triangle area representation (TAR) descriptor for non-rigid shapes which normalize the continuous contour using equidistant vertices normalization and discover the convexity/concavity of each point at different scales. The dynamic space warping (DSW) algorithm is used for similarity matching. It satisfies RST invariance property and it supports certain degree of occlusion and noise.

A shape descriptor by Chahooki et al. [7] integrates contour-based and region-based information for manifold learning based shape retrieval using fusion of dissimilarity measures. Lionel et al. [11] integrates shape moments, principal component analysis for shape curve normalization to generate compact dual feature vector which consist of angles between successive points and radius signature i.e normal centroid distance. The Fourier descriptor is applied on them to describe the shape effectively. This shape descriptor is affine invariant and robust to distortions. Another one integrated approach called color-shape context by Diplaros et al. [8] combines color and shape information for illumination and viewpoint invariant object recognition. Their method is robust to poor illumination, viewpoint changes, partial occlusion and cluttering.

The common challenges present in existing shape descriptors are: viewpoint changes, occlusion, cluttered scenes, noise, and shape deformation. A compact centroid distance shape descriptor described in the next section solves the problem of viewpoint changes.

3 The CCD descriptor for shape retrieval

In this work an efficient shape descriptor is proposed for image retrieval task. Compact Centroid Distance (CCD) descriptor is a affine invariant shape descriptor for image retrieval and object recognition task. Contour of the object is initially extracted. The CCD features are extracted by normalizing the shape into 'N' equal segments. Normalised points on the boundary of the object is located based on equal part area. Distance between centroid to all normalized points in the contour of the shape forms 1-D CCD feature vector. The CCD shape retrieval system measures the correspondence between the shapes using correlation coefficient metric.

3.1 CCD Feature Extraction

Object contour is extracted using Moore-edge following algorithm [14]. The coordinates are represented as $\Gamma_{\mu+i}$, i∈{0,1,2,…m-1}, 'm' is number of boundary pixels of the closed contour. We identified normalized contour points in the boundary of the object represented by $P_{\mu+i,}$ i∈{0,1,2,…N-1 (m>N) as shown in figure 1(b) . Sector area based normalization is used to identify the normalised points. Initially total area (S) of the object is computed, number of normalised points are initialised as N. To locate the coordinated position of those N points,

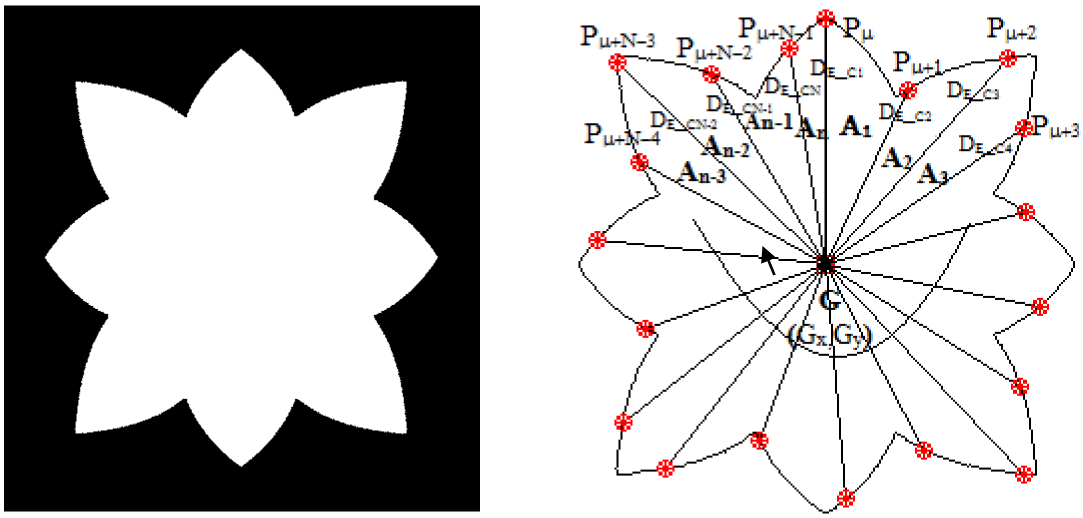
1. Initially compute the equal part area of a shape as S_{part}.

$$S_{part} = S/N \tag{1}$$

2. Take any two pixels along the boundary and compute the area enclosed between the pixels and the centroid as in eqn. 2.

$$SectorArea(P,\Gamma,\mathrm{G}) = \frac{1}{2} \times \left(\frac{r_P + r_\Gamma}{2} \right)^2 \times \theta \tag{2}$$

where (r_p, r_Γ) – distance from G to P and Γ respectively, and Θ - angle between r_p and r_Γ.



(a) (b)

Figure. 1. The CCD feature extraction process. (a) Input image (b) CCD feature extraction, the feature vector is (D_{E_C1}, D_{E_C2}, D_{E_C3}... D_{E_CN}).

r_p, r_Γ and Θ are computed as follows

$$r_p = ((P_x - G_x)^2 + (P_y - G_y)^2)^{1/2} \text{ and } r_\Gamma = ((\Gamma_x - G_x)^2 + (\Gamma_y - G_y)^2)^{1/2}$$

$$\theta = \tan^{-1}\left|\frac{M_P - M_\Gamma}{1 + M_P.M_\Gamma}\right|$$

where M_P, M_Γ - slope of lines (PG), and (ΓG).

$$M_P = \frac{P_y - G_y}{P_x - G_x},\ M_\Gamma = \frac{\Gamma_y - G_y}{\Gamma_x - G_x}$$

$\Gamma_{\mu+i}$ is incremented till the computed part area approximately approaches to S_{part}. When this criteria is satisfied then mark the pixel coordinates as $P_{\mu+1}$. This procedure is repeated to find all the remaining normalized points. The normalized points obtained after OAN are represented by (P_μ, $P_{\mu+1}$, $P_{\mu+2}$, ... $P_{\mu+N-1}$). In figure 1(b), the sector area between the points (P_μ, $P_{\mu+1}$, G) is A_1. Similarly the area between every adjacent normalized points and centroid forms equal part area vector {A_1, A_2, A_3,...A_{n-2}, A_{n-1}, A_n}.

The Euclidean distance as in equation (3) is used to find centroid to normalized boundary point distance value. And is represented as {D_{E_C1}, D_{E_C2}, D_{E_C3}... D_{E_CN}} in figure 1(b).

$$D_{E\_C}(G, P_\mu) = \sqrt{(G - P_\mu)^2} = |G - P_\mu| \qquad (3)$$

The performance of CCD shape descriptor is compared against normal centroid distance (CD) function. In CD feature extraction all boundary coordinates ($\Gamma_{\mu+i}$, and i∈ {0,1,2,…m-1}) of the shape is considered and its distance to centroid (G) are calculated. The features extracted in CD are denoted as {D_{E_N1}, D_{E_N2}, D_{E_N3}... D_{E_Nm}}. The size of CD feature vector is equals to total number of boundary pixels i.e., ‘m’.

3.2 Similarity Matching

The shape similarities are assessed by performing correlation coefficient on CCD features of the corresponding shapes. The correlation coefficient metric measures the association between two shapes P and Q which is determined by the following formula (4).

$$r(f_p, f_q) = \frac{\sum_{p,q=0}^{N-1}(f_p - \bar{f}_p)(f_q - \bar{f}_q)}{\sqrt{\left(\sum_{p=0}^{N-1}(f_p - \bar{f}_p)^2\right)\left(\sum_{q=0}^{N-1}(f_q - \bar{f}_q)^2\right)}} \tag{4}$$

Where (f_p, f_q) are CCD feature vectors of the shapes P and Q and its mean values are ($\bar{f}_p$, $\bar{f}_q$) respectively. The correlation between two feature vectors is in the range -1 to +1, where -1 is total negative correlation and +1 is total positive correlation. The intermediate values of correlation result shows degree of closeness between two shapes.

4 Results and Discussions

This section demonstrates the effectiveness of the CCD descriptor for affine invariant shape retrieval problems investigated on different testing conditions. The MPEG-7 CE Shape-1 Part-B dataset images are used for experimental purpose. The dataset is described by Latecki [10], which consists of 1400 images grouped into 70 classes and each class have 20 single shape object images. A subset of images from this dataset i.e., 60 numbers of different classes of shapes are chosen randomly and affine transformations like rotation, scaling and shearing are applied on each shape to generate 20 affine transformed images per shape. The affine transformed images generated by using the parameter values are shown in Table I.

Table I. The parameter values used to produce affine transformed images.

Transformation Name	Range	Interval	Number of affine transformed images
Rotation	1 to 360 degree	30 degree	12
Scaling	0.90 to 1.10	0.05	04
Shearing	-0.16 to +0.16	0.08	04
Total number of affine transformed images per class :			20

After affine transformation, a new dataset is formed which consists of 1200 (60x20) images. The experiments are conducted on the images in new dataset and the query image itself is present in this dataset.

During shape feature extraction, the normalization parameter N=64 is chosen that is common for all the query and dataset images. It partitions the shape into 64 equal part areas which produces 64 numbers of normalized contour points. The CCD calculated from these points produces a 1-D feature vector of size 64. In the case of CD descriptor, the feature vector size is proportional to number of boundary points of the shape.

The shape retrieval experiment is applied on 60 numbers of different classes of shapes present in the dataset. The performance comparison between CD and CCD descriptors on 25 different classes are shown in figure 2. The bar graph x-axis denotes image of 25 different shapes and the y-axis represents retrieval rate of CD and CCD on different images. This retrieval rate is calculated by the percentage of relevant (affine transformed) images retrieved in top 20 results. The CCD provides more retrieval rate i.e., double times better than CD in most of the classes. The average retrieval rate is calculated for both CD and CCD descriptors and is mentioned in the Table II. The average retrieval rate in equation (5) is calculated by the summation of retrieval percentages of all 60 classes divided by total number of classes i.e., 60.

$$Average\_Retrieval\_Rate = \left(\sum_{i=1}^{60} retrieval\_percentage_i\right)/60 \tag{5}$$

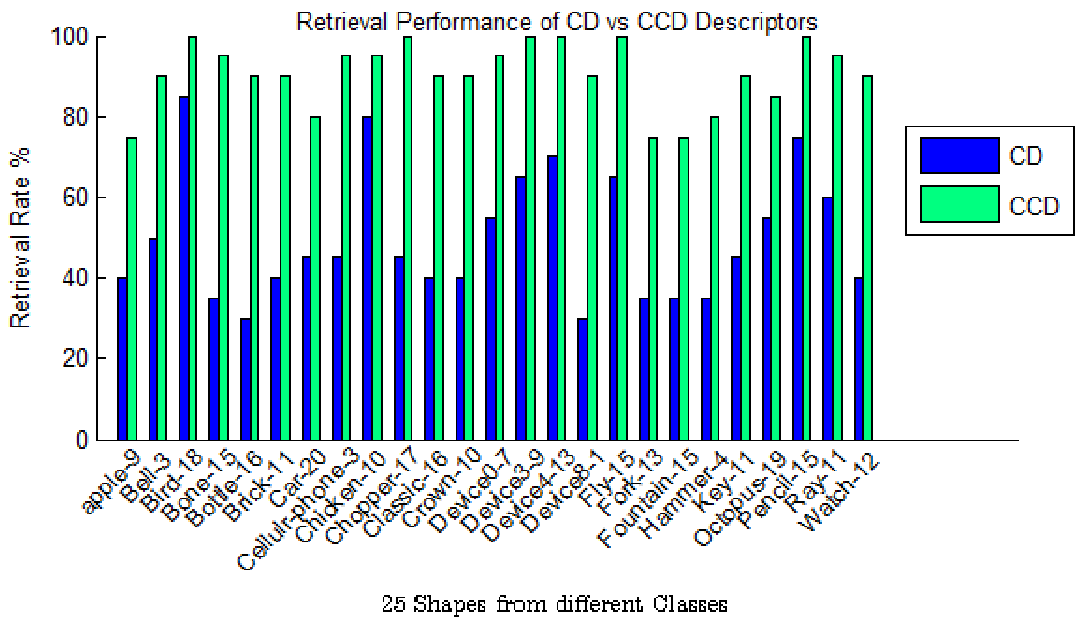


Figure. 2. The retrieval rate of CD vs. CCD on 25 different classes of shapes present in the dataset.

Table II. Average retrieval rate of CD and CCD descriptors.

Si. No	Method	Average Retrieval Rate
1	CD	61.75 %
2	CCD	86.00 %

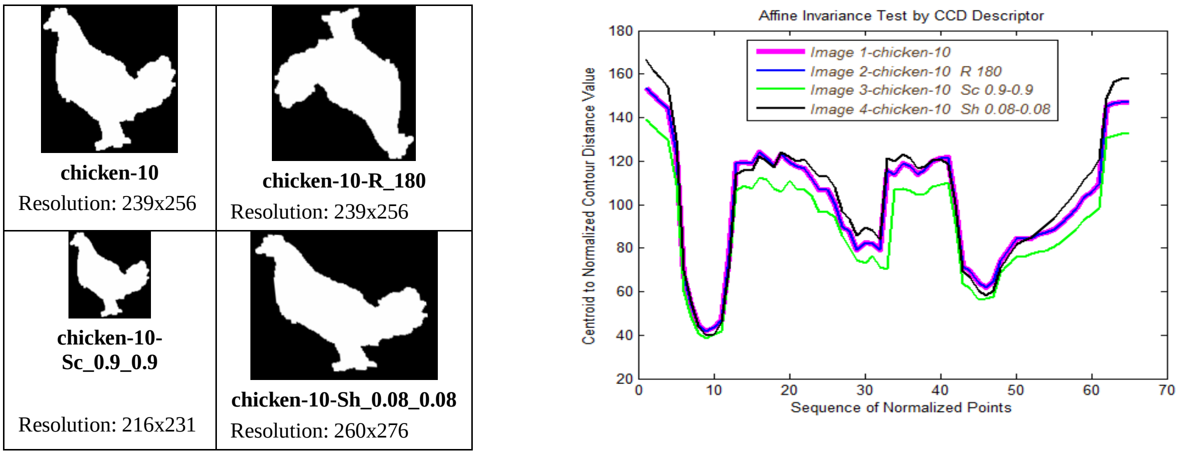
The retrieval results shown in Table II reveals that, the CCD shape descriptor is more robust to capture affine invariant shape information which is useful for identifying object shapes that undergoes different forms of geometrical transformations occurred due to viewpoint changes and it has good discrimination power to identify intra-class shapes.

The shape descriptor is also tested on another aspect i.e., the discrimination power of CCD descriptor is measured by two experiments. The first one is to test the ability of the descriptor on affine invariance property. The query image and its affine transformed versions are compared and the features extracted are plotted. The figure 3(a) shows the images used in the experiment and features are plotted in figure 3(b).

The second experiment of CCD descriptor is conducted on different classes of shapes present in the dataset. In figure 4(a) four different images used are shown and its features are plotted in figure 4(b). The variations in the curve imply that difference in the spatial distribution of pixels for dissimilar shapes.

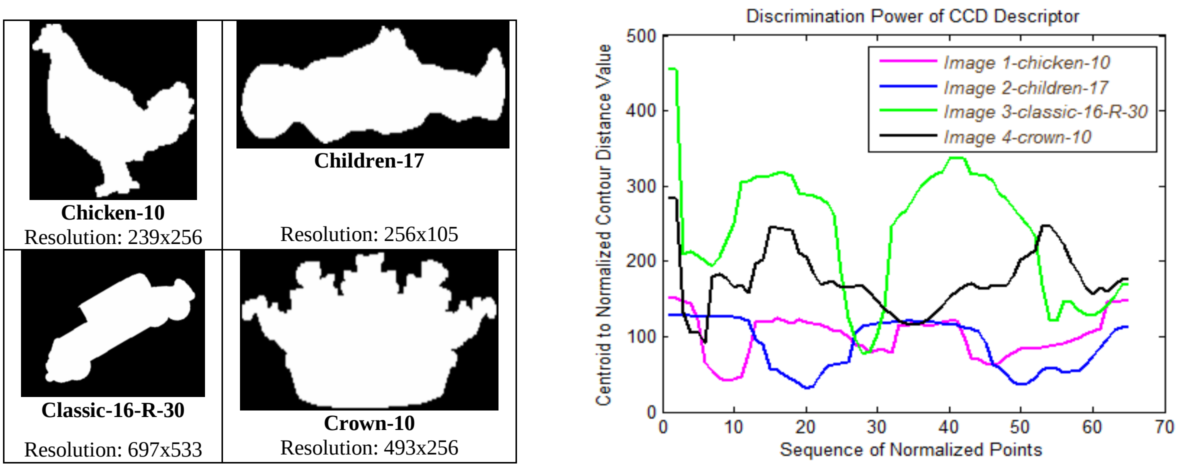
The matching time complexity is one of the important property that decides efficiency of the shape descriptors. The CD descriptor's matching time is huge because it accounts all contour points of the shape. In CCD descriptor the matching time complexity is directly proportional to number of part area segments or number of normalized contour points. The matching time complexities of CD and CCD are O (m) and O(N) respectively, m>N.

Apart from effectiveness in retrieval, the CCD descriptor is more compact and its matching time complexity is very small depends upon normalization value. The limitations in this descriptor are starting point variation affects retrieval of shear transformed images and to fix standard 'N' value. The feature work directions are to fix starting point problem and select best possible 'N' value.



(a) (b)

Figure. 3. The CCD descriptor is plotted to prove affine invariance property. Four curves plotted from the query image, and its affine transformed (rotate, scale and shear) images.



(a) (b)

Figure. 4. Illustration of discrimination power of CCD descriptor on different object shapes.

5 Conclusions

In this paper, a simple, fast and more compact contour based affine invariant shape descriptor has been presented for CBIR applications. The sector area based object area normalization was performed on closed contour object shapes which partitioned the object into 'N' number of equal part area segments with respect to centroid. The centroid distance values calculated from all the normalized contour points provide shape information in the form of 1-D feature vector of size ‘N’. For the experimental investigations, a new dataset created from the MPEG-7 CE Shape-1 Part-B dataset images. The affine transformed versions of intra-class and inter-class shapes are generated by appropriate affine parameters. The correspondence between query and affine transformed versions of images found in the dataset are determined by the correlation coefficient metric accordingly the images are retrieved. The experimental results show that, the CCD descriptor using object area normalization works better than basic CD function. The matching complexity of CCD is very less than CD descriptor. The optimal selection of number of normalized contour points ‘N’ is one among the feature enhancement work to standardize the selection of ‘N’.

References

[1] S. Abbasi, F. Mokhtarian and J. Kittler, Curvature Scale Space image in Shape Similarity Retrieval, *Multimedia Systems-Springer Varlag* 1999, 467-476. http://dx.doi.org/10.1007/s005300050147

[2] N. Alajlan, I. E. Rube, M. S. Kamel and G. Freeman, Shape retrieval using triangle-area representation and dynamic space warping, *Pattern Recognition* 40 (2007), 1911 – 1920. http://dx.doi.org/10.1016/j.patcog.2006.12.005

[3] A. Andreopoulos and J. K. Tsotsos, A Computational Learning Theory of Active Object Recognition Under Uncertainty, *International Journal of Computer Vision (IJCV), Springer,* 2013, 95-142. http://dx.doi.org/10.1007/s11263-012-0551-6

[4] N. Arica and F. Vural, BAS: a perceptual shape descriptor based on the beam angle statistics, *Pattern Recognition Letters*, Volume 24, 2003. http://dx.doi.org/10.1016/s0167-8655(03)00002-3

[5] P. Arjun, T. T. Mirnalinee, S. Sindhuja and G. Bharathiraja, Affine Invariant Shape Descriptor using Object Area Normalization, *International Conference on Artificial Intelligence and Evolutionary Algorithms in Engineering Systems (ICAEES)*, 2014. http://dx.doi.org/10.1007/978-81-322-2119-7_115

[6] J. Ashley, R. Barber, M. Flickner, J. Hafner, D. Lee,W. Niblack and D. Petkovic, Automatic and semiautomatic methods for image annotation and retrieval in QBIC, *Proc. of Storage and Retrieval for Image and Video Databases (SPIE) III*, 2420(1994), 24–35. http://dx.doi.org/10.1117/12.205303

[7] M. A. Z. Chahooki and N. M. Charkari, Shape retrieval based on manifold learning by fusion of dissimilarity measures, *IET Image Processing*, Vol. 6, Issue. 4(2012), 327–336. http://dx.doi.org/10.1049/iet-ipr.2010.0548

[8] A. Diplaros, T. Gevers and I. Patras, Combining Color and Shape Information for Illumination-Viewpoint Invariant Object Recognition, *IEEE Transactions on Image Processing*, Vol. 15, No. 1(2006), 1-11. http://dx.doi.org/10.1109/tip.2005.860320

[9] J. Eakins, Content-based Image Retrieval, *The JISC Technology Applications Programme Report 39*, 1999.

[10] L.J. Latecki, R. Lakamper and U. Eckhardt Shape Descriptors for Non-rigid Shapes with a Single Closed Contour, *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2000, 424-429. http://dx.doi.org/10.1109/cvpr.2000.855850

[11] P. Lionel Evina Ekombo, N. Ennahnahi, M. Oumsis and M. Meknassi, Application of affine invariant Fourier descriptor to shape based image retrieval, *International Journal of Computer Science and Network Security*, VOL.9 No.7, July 2009, 240-247.

[12] Y. Liu, D. Zhang, G. Lu and W.Y. Ma, A survey of content-based image retrieval with high-level semantics, *Pattern Recognition*, 2006, 262 – 282. http://dx.doi.org/10.1016/j.patcog.2006.04.045

[13] Y. Mingqiang, K. Kidiyo and R. Joseph, Affine Invariance Contour Descriptor Based on the Equal Area Normalization, *IAENG International Journal of Applied Mathematics*, 36:2, 2007.

[14] The Moore Edge-following algorithm, Available at: http://www.imageprocessingplace.com/downloads_V3/root_downloads/tutorials/contour_tracing_Abeer_George_Ghuneim/moore.html.

Received: December 15, 2014; Published: March 21, 2015