Introducing an Adaptive Kernel Density Feature Points Estimator for Image Representation

Tranos Zuva and [2]Oludayo, O. Olugbara

[1,4]Department of Computer System Engineering

Tshwane University of Technology,
Pretoria, South Africa,
[2]Department of Information Technology

Durban University of Technology

Durban, South Africa,

[3]Sunday, O. Ojo and [4]Seleman, M. Ngwira
[3]Faculty of Information and Communication Technology

Tshwane University of Technology

Pretoria, South Africa, 0001

Abstract :

This paper provides an image shape representation technique known as Adaptive Kernel Density Feature Points Estimator (AKDFPE). In this method, the density of feature points within defined rings (bandwidth) around the centroid of the image is obtained in the form of a vector. The AKDFPE is then applied to the vector of the image. AKDFPE is invariant to translation, scale and rotation. This method of image representation shows improved retrieval rate when compared to Kernel Density Feature Points Estimator (KDFPE) method. Analytic analysis is done to justify our method, which was compared with the KDFPE to prove its robustness.

Keywords *- Kernel Density Function, Similarity, Image Representation, Segmentation, Density Histogram*

I. INTRODUCTION

The huge collection of digital images on personal computers, institutional computers and Internet necessitates the need to find a particular image or a collection of images of interest. This has motivated many researchers to find efficient, effective and accurate algorithms that are domain independent for representation, description and retrieval of images of interest. There have been many algorithms developed to represent, describe and retrieve images using their visual features such as shape, colour and texture [1], [2], [3], [4], [5-7]. The visual feature representation and description play an important role in image classification, recognition and retrieval. A successful image representation and description is dependent on the selection of suitable image features to encode and quantify these features [4].

Shape representation and description have been dominant in research area of image processing because shape is considered to be the basis of human visual recognition [4]. The shape representation can be classified as region based or contour based. The contour based techniques use the boundary of shape to describe an object. It is commonly believed that human beings can differentiate objects by their boundaries or contours [2]. Usually, most objects form shapes with defined contours, making the use of these techniques most appealing. The techniques can generally be applied to different application areas with a considerable success. The techniques have a low computation complexity as compared to region based techniques and they are sensitive to noise. These techniques in this group are well described in [5].

The region based shape representation uses the boundary pixels and the interior pixels of the shape. This group of shape representation algorithms are robust to noise, shape distortion and they are applicable to generic shapes [8]. These techniques can be found in [5]. This paper proposes Adaptive Kernel Density Feature Points Estimator (AKDFPE) image shape representation technique. This method imitates human visualization of image object shape and matching similar object shapes. A comparison of retrieval of similar image object shapes is done between AKDFPE and Kernel Density Feature Points Estimator (KDFPE) [7] representation of image object shapes.

II. SHAPE REPRESENTATION BY AKDFPE

This method describes the feature points within rings in an image grid. Assume we have a silhouette object shape segmented by some means such as active contour without edges [9] and let the feature points set $P(x, y)$ (intensity function) of the object shape be defined as:

$$P(x, y) = p_i(x, y), i = 1, 2, ...n, n \varepsilon \mathrm{N} \quad (1)$$

The centroid of the object shape is calculated. The following formulae will be used to calculate the centroid [10],[11]:

$$x_c = \frac{m_{1,0}}{m_{0,0}} \quad (2)$$

$$y_c = \frac{m_{0,1}}{m_{0,0}} \quad (3)$$

where $m_{1,0}, m_{0,1}, m_{0,0}$ are derived from the silhouette moments given by

$$m_{i,j} = \sum_x \sum_y x^i y^j P(x, y). \quad (4)$$

The theorems that guarantee the uniqueness and existence of silhouette moments can be found in [10]. For silhouette image $P(x, y)$, $m_{0,0}$ the moment of zero order represents the geometrical area of the image region and $m_{1,0}, m_{0,1}$ moment of first order represents the intensity moment about the $y - axis$ and $x - axis$ of the image respectively. The centroid (x_c, y_c) gives the geometrical centre of the image region.

Suppose the size of the grid occupied by the object shape is $N \times N$. The vector dimension to represent the density of object shape will be $N - 1$. From the centroid we count the number of image pixels in the rings with defined equal width around the centroid. The number of image pixels in each ring is given as $X_i = (n_1, n_2, ...n_m)$ where m is the number of rings from the centroid. The Adaptive Kernel Density Feature Points Estimator (AKDFPE) is applied. The AKDFPE using the modified Loftsgaarden-Quesenberry nearest-neighbour kernel is given in equation (5).

$$\hat{f}(x) = \frac{1}{m h_{k_c}(x)} \sum_{i=1}^{m} K\left(\frac{x_1^m - x}{h_{k_c}(x)}\right) \quad (5)$$

The number of nearest neighbours k_c controls the level of smoothing of clusters c, $i = 1,2,3..........,m$. $K(\bullet)$ is the kernel function, m is the number of rings and h_{k_c} is the bandwidth per cluster. We calculate the optimal bandwidth h_{oc_j} per cluster. Then we recalculate the vector elements of the image, using equation (6) that follows:

$$f(x_i) = \frac{1}{h_{oc_j}} K\left(\frac{x_i^m - \bar{x}}{h_{oc_j}}\right) \tag{6}$$

where $j = 1,2,3,.....n$.

The optimal bandwidth h_{oc_j} for each cluster in image shape vector is calculated using the following second order Gaussian plug-in formula [12]:

$$h_{oc_j} = 1.059 m_j^{\frac{-1}{5}} s_j \tag{7}$$

where m_j is the number of rings in a cluster and s_j is the cluster standard deviation. The clusters are formed through the percentage of the image occupying the ring. For the purpose of illustration of this method, suppose object shape features are given on a grid as shown in Fig 1.

0.0	**1,0***	**2,0***	**3,0***	**4,0***
0,1	**1,1***	**2,1***	**3,1***	**4,1***
0,2*	**1,2***	**2,2***	**(x_c, y_c) 3,2**	**4,2***
0,3	1,3	2,3	**3,3***	**4,3***
0,4	1,4	2,4	**3,4***	**4,4***

Fig. 1 Segmented Object Shape

The star red-bolded indicate the "on" pixels, which belong to the image. The size of the grid occupied by the object shape is 5×5. The centroid calculated using equations (2) and (3) is (3,2), the centroid pixel is in blue. The first rectangle boundary in Fig.1 is made up of the following pixels (2,1), (3,1), (4,1), (4,2), (4,3), (3,3), (2,3), (2,2) and there are seven "on" pixels that constitute our first element of the vector. The preliminary vector representation of object shape in Fig.1 is (7, 8, 1). The vector above will be represented as follows in the standardized way:

$x_1^3 = (28,16,1)$

and percentages are as follows

$\%s = (88,50,3)$

It means they belong to three different clusters. In this case we calculate global bandwidth. From now we apply the Adaptive Kernel Density Feature points Estimator (AKDFPE).

The AKDE is given as

$$\hat{f}_{h(x)}(x) = \frac{1}{3}\sum_{i=1}^{3} K_{h(x)}(x - x_i) = \frac{1}{3h(x)}\sum_{i=1}^{3} K\left(\frac{x - x_i}{h(x)}\right) \quad (8)$$

We then calculate the optimal bandwidth h_{oc_j} for each cluster in an image shape vector. Then we recalculate the vector elements of the image, using the univariate balloon estimator using modified Loftsgaarden-Quesenberr k^{th} nearest neighbourhood given in equation (9).

$$\hat{f}_B(x) = \frac{1}{nh(x)}\sum_{i=1}^{n} K\left(\frac{x_i - x}{h(x)}\right) \quad (9)$$

The estimate of $\hat{f}_B(x)$ is an average of identically scaled kernels centred at each data point.

$$f_1^3(x) = \frac{1}{h_{oc_1}} K\left(\frac{x_1^3 - \bar{x}}{h_{oc_1}}\right) \quad (10)$$

where $K(\bullet)$ is the kernel function.

This is how the images will be represented.

III. SIMILARITY MEASUREMENT

In order to measure the similarity of the images we used the cosine coefficient given in [13] as

$$s_{\cos} = \frac{\sum_{i=1}^{m} P_i Q_i}{\sqrt{\sum_{i=1}^{m} P_i^2}\sqrt{\sum_{i=1}^{m} Q_i^2}} \quad (11)$$

The cosine coefficient, which is also called angular metric, is the normalized inner product of two vectors because it measures the angle between those vectors. The cosine coefficient has lower and upper bounds of 0 and 1 respectively. This makes it more suitable than Euclidean metric to establish comparison of results produced by two different image retrieval methods such as KDFPE and AKDFPE.

Iv . ACCURACY MEASUREMENT

The accuracy of an image retrieval system is generally measured by calculating recall, precision and effectiveness of the system. The following formulas were used [1]:

$$recall = \frac{A}{N} \quad (12)$$

$$precision = \frac{A}{A+C} \qquad (13)$$

$$effectiveness = \begin{cases} \frac{A}{N} & if\ T \rangle N \\ \frac{A}{T} & if\ T \le N \end{cases} \qquad (14)$$

In this case, A is the number of relevant image objects retrieved, B is the number of relevant image objects not retrieved, C is the number of irrelevant image objects retrieved, N total relevant in the database and T is the user required number of relevant image retrieval.

V. EXPERIMENTATION

The main objective of the experimentation is to find the effectiveness of AKDFPE method and to compare it with other representation methods. In this case a comparison is made with the KDFPE. The cosine coefficient similarity measure is used in retrieval of similar image objects. An image database of 200 shop items shapes is created. Some of the image objects are rotated at 90, 180 and 270 degrees. The images that are rotated were not rotated lossless, meaning degradation of the image object occurred during rotation. The image objects were of different dimensions $M \times N$ or $N \times N$ where M and N are real numbers when they are brought to the system. The images that are used only have one image object with a homogeneous background. The image object shape of grid dimension 45×45 is segmented using the Chan and Vese active contour without edge [7]. All images are converted to gray scale images. They are then represented using AKDFPE and KDFPE. Each image was used as a query and the retrieval rate was measured using the Bull's Eye Performance (BEP), recall and precision performance. Matlab 7.6 was used to implement the system. Example of classes of shapes experimented with are given in Fig 2. In each class there are ten elements with some items rotated and scaled.

Fig. 2 Examples of Classes of Items in Database

VI. RESULTS

Different types of televisions are in the collection of televisions in the database [13]. Some of the segmented televisions are shown in Fig.3 below. The AKDFPE is capable of matching most of televisions despite errors in segmentation, noise and distortion of the television shapes because of transformations such as scaling and rotation.

Proceedings is available @ International Journals of Information Technology and Computer Science (IJITCS) (ISSN No : 2091-1610) (http://www.ijitcs.com) , International Journal of Wireless Information Networks & Business Information System (http://www.e-winbis.com) (ISSN No : 2091-0266)

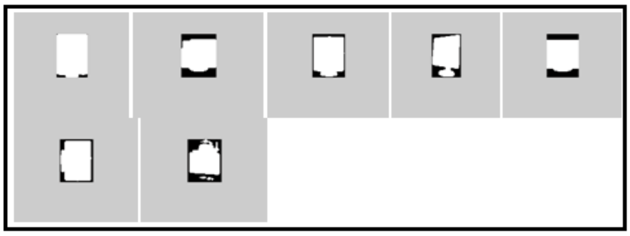

Fig. 3 Segmented shapes that were considered similar by AKDFPE

Fig. 4 shows the images considered similar to the query image on the top left of the figure. It can be seen that the query image is part of the retrieved images, indicating that it belongs to the database. In this sample of retrieval in Fig. 4, AKDFPE has a 90% precision while KDFPE has a 80% precision.

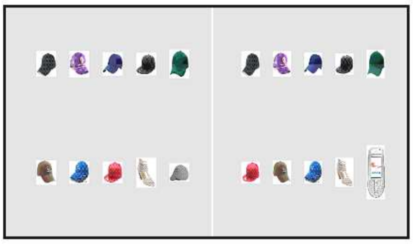

Fig. 4 Ten retrieval results of AKDFPE on the left and KDFPE on the right

The result in Fig. 5 shows that AKDFPE is better in retrieving images that are in the database using query images that belong to the database. It can be appreciated that the difference is not very substantial when it comes to querying the database with images in it. The Bull's Eye Performance (BEP) of KDFPE is 92.40% while AKDFPE is 93.76%.

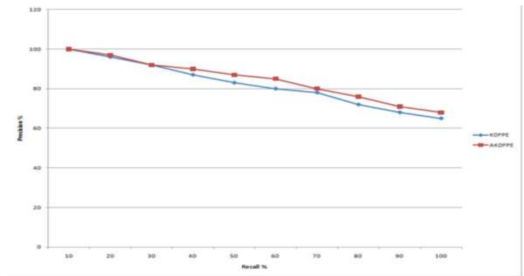

Fig. 5 Ten retrieval results of AKDFPE on the left and KDFPE on the right

VII. CONCLUSION

From the results it can be concluded that AKDFPE method of image object representation is able to differentiate similar object shapes just as human beings perceive image object shapes. The ability to calculate the variable optimal bandwidth per given image seems to give AKDFPE an advantage over KDFPE method. The KDFPE calculates the global optimal bandwidth which does not take into consideration the areas that are sparsely or densely populated by the image pixels. Different kernel functions are going to be used in future to investigate if they have any effect in the retrieval rate. AKDFPE is capable of overcoming errors in segmentation and is robust to segmentation noise. In general both methods show high retrieval rate showing that usage of proper estimation techniques in image retrieval might bring solution to finding generic image representation methods.

REFERENCES

1. M. X. Ribeiro, et al., "Statistical Association Rules and Relevance Feedback: Power Allies to Improve the Retrieval of Medical Images," Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems, 2006.
2. D. Zhang and G. Lu, "Review of shape representation and description techniques," Pattern Recognition Society, vol. 37, pp. 1-19, 2004.
3. Y. Li and L. Guan, "An effective shape descriptor for the retrieval of natural image collections," in. Proceedings of the IEEE CCECE/CCGEI, Ottawa, 2006, pp. 1960-1963.
4. X. Zheng, et al., "Perceptual shape-based natural image representation and retrieval," in Proceedings of the IEEE International Conference on Semantic Computing, 2007, pp. 622-629.
5. Y. Mingqiang, et al., "A survey of shape feature extraction techniques," in Pattern Recognition, 2008, pp. 43-90.
6. D. C. Tran and K. Ono, "Content-based image retrieval: Object representation by the Density of feature Points," pp. 213-218, 2000.
7. T. Zuva, et al., "Object Shape Representation by Kernel Density Feature Points Estimator," in First International Workshop on Signal and Image Processing (SIP), Bangalore, India, 2012, pp. 209-216.
8. E. M. Celebi and A. Y. Aslandogan, "A comparative Study of Three Moment-Based Shape Descriptors," Proceedings of the International Conference on Information Technology: Coding and Computing, 2005.
9. T. F. Chan and L. A. Vese, "Active Contours Without Edges," IEEE, vol. 10, pp. 266-277, 2001.
10. J. Flusser, et al., Moments and moment invariants in pattern recognition. West Sussex: John Wiley & Sons Ltd., 2009.
11. R. Mukundan and K. R. Ramakrishnan, Moment functions in image analysis: theory and applications. Singapore: World Scientic Publishing Co. Pte. Ltd., 1998.
12. J. S. Simonoff, "Smoothing Methods in Statistics," in Springer Series in Statistics, Springer, Ed., ed, 1996.
13. S.-H. Cha, "Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions," International Journal of Mathematical Models and Methods in Applied Sciences, vol. 1, pp. 300-307, 2007.