Object Shape Representation by Kernel Density Feature Points Estimator

Tranos Zuva[1], Oludayo O. Olugbara[2] , Sunday O. Ojo[3] and Seleman M. Ngwira[4]

[1, 4]Department of Computer Engineering, Tshwane University of Technology, Pretoria, South Africa
zuvat@tut.ac.za
[3]Faculty of Information and Communication Technology,
[2]Department of Information Technology, Durban University of Technology, Durban, South Africa

ABSTRACT

This paper introduces an object shape representation using Kernel Density Feature Points Estimator (KDFPE). In this method we obtain the density of feature points within defined rings around the centroid of the image. The Kernel Density Feature Points Estimator is then applied to the vector of the image. KDFPE is invariant to translation, scale and rotation. This method of image representation shows improved retrieval rate when compared to Density Histogram Feature Points (DHFP) method. Analytic analysis is done to justify our method and we compared our results with object shape representation by the Density Histogram of Feature Points (DHFP) to prove its robustness.

KEYWORDS

Kernel Density Function, Similarity, Image Representation, Segmentation, Density Histogram

1. INTRODUCTION

With vast collection of digital images on personal computers, institutional computers and the Internet, the need to find a particular image or a collection of images of interest has increased tremendously. This has motivated the 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 that have been developed to represent, describe and retrieve images using their visual features such as shape, colour and texture [1], [2], [3], [4], [5]. Visual feature representation and/or 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 quantification of 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.

Natarajan Meghanathan, et al. (Eds): ITCS, SIP, JSE-2012, CS & IT 04, pp. 209–216, 2012.

DOI : 10.5121/csit.2012.2119

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 (distribution of all pixels within contour). This group of shape representation algorithms are robust to noise, shape distortion and they are applicable to generic shapes [6]. These techniques can be found in [5].

This paper proposes a Kernel Density Feature Points Estimator representation of an image object. 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 Kernel Density Feature Points Estimator and Density Histogram of Feature Points representation of image object shapes.

2. SHAPE REPRESENTATION BY KDFPE

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 [7] and let the feature points set $P(x, y)$ (intensity function) of the object shape be defined as

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

The centroid of the object shape is calculated. The following formulae will be used to calculate the centroid [8],[9]:

$$x_c = \frac{m_{1,0}}{m_{0,0}} \tag{2}$$

$$y_c = \frac{m_{0,1}}{m_{0,0}} \tag{3}$$

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

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

The theorems that guarantee the uniqueness and existence of silhouette moments can be found in [8]. 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 NxN. 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 Kernel Density Feature Points Estimator (KDFPE) is then applied. The Second-order Gaussian Kernel Density Estimator (SGKDE) is used.

The SGKDE is given in [10] as

$$f(x) = \frac{1}{mh} \sum_{i=1}^{m} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{x - X_i}{h}\right)^2} \tag{5}$$

The optimal bandwidth h_o for each image shape is calculated using the following second order Gaussian plug-In formula [11]:

$$h_o = 1.059 m^{\frac{-1}{5}} s \tag{6}$$

where m is the number of rings from the centroid and s is sample standard deviation.
The vector elements of the image are recalculated using the following:

$$f(x_i) = \frac{1}{h_o \sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{X_i}{h_o}\right)^2} \tag{7}$$

$f(x_i)$ becomes the image representation vector.

2.1. Illustration

Suppose object shape features are given on a grid as shown in figure 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**	**3,2**	**4,2**
0,3	1,3	2,3	**3,3**	**4,3**
0,4	1,4	2,4	**3,4**	**4,4**

Figure 1**.** Segmented object shape

The red-bold indicate the “on” pixels (pixels that belong to the image). The size of the grid occupied by the object shape is 5X5. The centroid calculated by the two formulas above (2) and (3) is (3, 2), the centroid pixel is in blue. The first rectangle boundary in figure 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 figure 1 is

$$(7, 8, 1)$$

The Kernel Density Estimator (KDE) is then applied. As mentioned earlier on the second-order Gaussian KDE is going to be used. The sample points (in our example) with the unknown distribution function are:

$$(7,8,1)$$

The Second-order Gaussian KDE is given as:

$$f(x) = \frac{1}{3h}\sum_{i=1}^{3}\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}\left(\frac{x-X_i}{h}\right)^2} \tag{8}$$

The optimal bandwidth h_o for each image shape is calculated. Then we the recalculate the vector elements of the image, using the following

$$f(x_i) = \frac{1}{h_o\sqrt{2\pi}}e^{\frac{-1}{2}\left(\frac{X_i}{h_o}\right)^2} \tag{9}$$

This is how the images will be represented.

3. SIMILARITY MEASUREMENT

To measure the similarity of the images we used the Cosine Coefficient given in [12] as

$$s_{\cos} = \frac{\sum_{i=1}^{m} P_i Q_i}{\sqrt{\sum_{i=1}^{m} P_i^2}\sqrt{\sum_{i=1}^{m} Q_i^2}} \tag{10}$$

The equation above is the normalized inner product and called the cosine coefficient because it measures the angle between two vectors. It is often called angular metric.

4. ACCURACY MEASUREMENT

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

$$recall = \frac{A}{N} = \frac{total\ relevent\ in\ the\ query\ results}{total\ relevent\ in\ database} \tag{11}$$

$$precision = \frac{A}{A+C} = \frac{total\ relevent\ in\ the\ query\ results}{total\ in\ the\ query\ results} \tag{12}$$

$$effectiveness = \begin{cases} \frac{A}{N} \; if \; T \rangle N \\ \frac{A}{T} \; if \; T \leq N \end{cases} \tag{13}$$

Where 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 and T is the user required number of relevant image retrieval.

5. EXPERIMENTATION

The main objective is to find the effectiveness of KDFPE method and also compare it with other representation methods. In this case a comparison is made with Density Histogram Feature Points. 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 MXN or NXN where M and N belong to 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 a 45X45 grid is segmented. All images are converted to gray scale images. They are then represented using KDFPE and DHFP. The average precision is calculated for retrieval of images in the database using queries of images from the database. The average effectiveness is calculated for retrieval images in the database using queries of images captured using camera enabled devices. Matlab 7.6 was used to implement the system. Example of classes of shapes experimented with are given in figure 2.

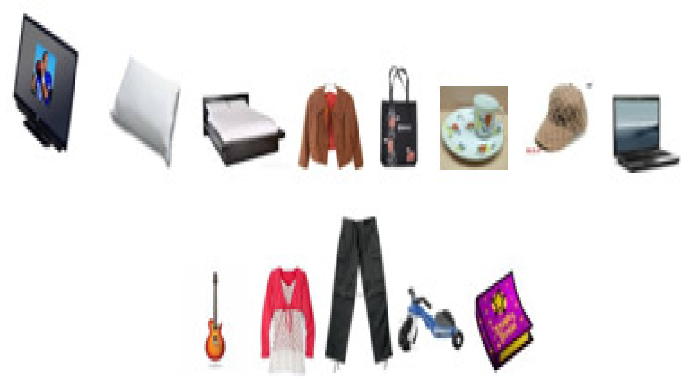

Figure 2. Examples of classes of items in database

In each class there are ten elements with some items rotated and scaled.

6. RESULTS

6.1. Similarity Matching

Different types of televisions are in the collection of televisions in the database. Some of the segmented televisions are shown in figure 3 below. The KDFPE is capable of matching most of televisions despite errors in segmentation, noise and distortion of the television shapes due to transformations (scaling and rotation).

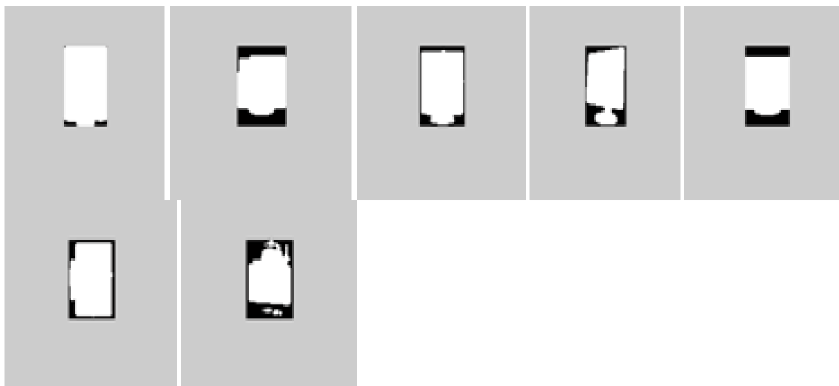

Figure 3. Segmented shapes that were considered similar by KDFPE

6.2. Comparison of Representation Methods (KDFPE and DHFP)

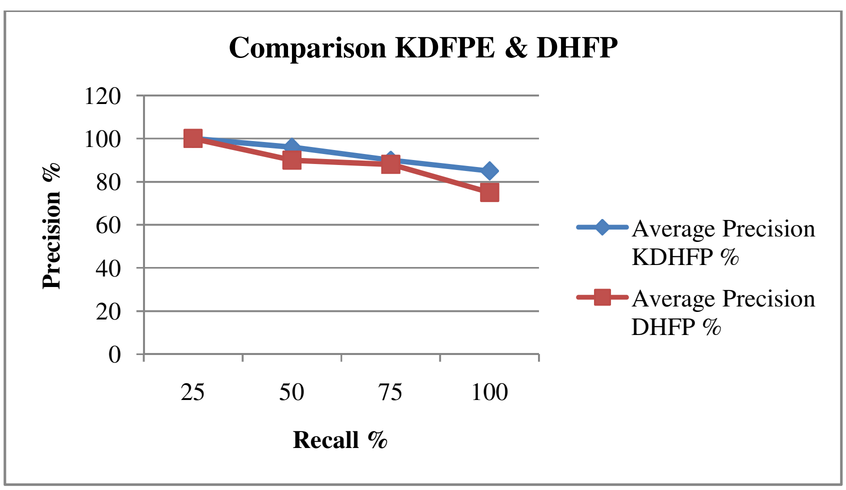

Figure 4. Comparison of KDFPE and DHFP

The results in figure 4 show that KDFPE 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.

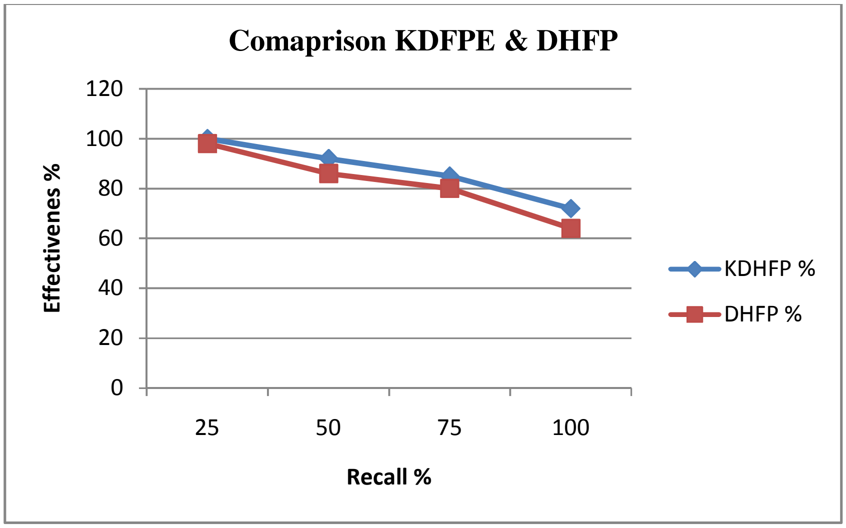

Figure 5. Comparison of effectiveness of the two methods

The results shown on figure 5 shows that KDFPE is better than the DHFP method when using a query that does not belong to the database. The figure 5 shows a very substantial difference in the effectiveness of the methods.

7. SUMMARY AND CONCLUSION

From the results it can be concluded that KDFPE method of image object representation is able to differentiate similar object shapes just as human beings perceive image objects shapes. The ability to calculate the optimal width seems to give KDFPE an advantage over DHFP method. The KDFPE method shows that it is capable of retrieving images similar to images captured by a camera enabled device as shown in figure 5. This characteristic is good in recommender systems because people are looking for images similar to their own captured images but not necessarily in the database. In future we are going to use different methods of calculating optimal width to investigate if we may improve the effectiveness of KDFPE. Different kernel functions are going to be used in future to investigate if they have any effect in the retrieval rate. It is important to have a retrieval system that is capable of matching images that are in database with images that are not in the database. KDFPE is capable of overcoming errors in segmentation and is robust to segmentation noise.

REFERENCES

[1] M. X. Ribeiro, et al.,(2006) "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.

[2] D. Zhang and G. Lu, (2004) "Review of shape representation and description techniques," Pattern Recognition Society, vol. 37, pp. 1-19.

[3] Y. Li and L. Guan, (2006) "An effective shape descriptor for the retrieval of natural image collections," in. Proceedings of the IEEE CCECE/CCGEI, Ottawa, pp. 1960-1963.

[4] X. Zheng, et al., (2007) "Perceptual shape-based natural image representation and retrieval," in Proceedings of the IEEE International Conference on Semantic Computing, pp. 622-629.

[5] Y. Mingqiang, et al., (2008) "A survey of shape feature extraction techniques," in Pattern Recognition, pp. 43-90.

[6] E. M. Celebi and A. Y. Aslandogan, (2005) "A comparative Study of Three Moment-Based Shape Descriptors," Proceedings of the International Conference on Information Technology: Coding and Computing.

[7] T. F. Chan and L. A. Vese, (2001) "Active Contours Without Edges," IEEE, vol. 10, pp. 266-277.

[8] J. Flusser, et al., (2009) "Moments and moment invariants in pattern recognition." West Sussex: John Wiley & Sons Ltd.

[9] R. Mukundan and K. R. Ramakrishnan, (1998) "Moment functions in image analysis: theory and applications." Singapore: World Scientic Publishing Co. Pte. Ltd.

[10] H. Shimazaki and S. Shinomoto, (2010) "Kernel bandwidth optimization in spike rate estimation," J Comput Neurosci, vol. 29, pp. 171-182.

[11] J. S. Simonoff, (1996) "Smoothing Methods in Statistics," in Springer Series in Statistics, Springer, Ed., ed.

[12] S.-H. Cha, (2007) "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.