Volume 4, Issue 7, July 2014 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

An Efficient and Easy Approach in Partial Shape Matching for Gesture Recognition

Miss S. A. Bhosale*
Department of CS & IT,
Dr. B. A. M. U,
Aurangabad, India

Mr. P. P. Chouthmal
Department of CS & IT,
Dr. B. A. M. U,
Aurangabad, India

Dr. K. V. Kale
Head, Department of CS & IT,
Dr. B. A. M. U,
Aurangabad, India

Abstract— ***We present a new approach for partial shape matching efficiently and more easily by using Local Tetra Pattern (LTrP) for feature extraction and Scale Invariant Feature Transformation (SIFT) algorithm for matching of feature points. The key idea and main contribution of this work lies in matching the textures and recognizing it accordingly by providing some special features like scale invariance, deformation tolerance and orientation free nature. At the heart of the proposed scheme lies a new shape descriptor that also permits the quantification of local scale. The shape descriptors are computed along open or closed contours in a spatially non-uniform manner. The LTrP texture pattern is proposed to extract the image texture features and then SIFT algorithm which is a powerful one for performing texture classification and recognition. Due to the properties of newly proposed shape descriptor, extraction method used and matching technique, the proposed approach performs partial shape matching in a more efficient way.***

Keywords— Gesture recognition, Local Tetra Pattern (LTrP), Partial shape matching, Shape descriptor, Scale Invariant Feature Transformation (SIFT)

I. INTRODUCTION

Shape matching is one of the fundamental problems in computer vision and digital image processing. It deals with finding the number of basic features in an image and by using those features calculating there similarities with other images to classify them appropriately.

The quality of the shape matching process depends on whether its final outcome agrees with human judgment. Shape matching is a very challenging problem. Shapes to be matched are typically the result of some kind of segmentation process which, being imperfect, may introduce a considerable amount of noise that needs to be handled and removed. In most of the situations, arbitrary differences in scale and orientation should not affect the matching method. Due to point of view dependencies, shape articulations and deformations in the regions, various 2D image projections of the shape of the same 3D object may vary considerably [1].

Normally the performance of shape matching techniques gets decreased due to various reasons like noise in images, large amount of deformation present, segmentation errors etc. Shape matching is a vast problem area in computer vision and has various applications but specifically such a system can be used to control any device without touching it as like Samsung smart TV or even in virtual systems also. Other applications may include medical imaging, CAD-CAM designs, security applications like face recognition, finger print recognition, shape classification etc.

In the current context of work, we are interested in addressing the two-dimensional shape matching problem by taking into consideration an existence of all the complicating things. Shapes are represented as binary images depicting foreground objects over their background. The Fig. 1 represents the overall partial shape matching concept and classifies query image of palm structure to an appropriate class based on the image features. Here three classes have been represented to classify a hand gesture image to an appropriate class according to the contour of it and by using different features.

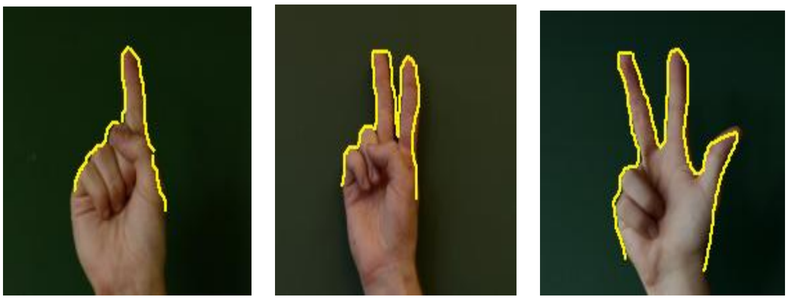

Fig. 1 Basic prototype images which represents (a) Class 'one', (b) Class 'two', (c) Class 'three'

II. LITERATURE SURVEY

purpose of good recognition rate a well-organized hand model is the basic need. T. B. Sebastian in [2] adopts recognition framework which is based on matching shock graphs of 2D shape outlines, where the distance between two shapes is defined to be the cost of the least action path deforming one shape to another. But still two demerits are their particularly the algorithm typically takes about 8 minutes on an SGI Indigo II and it is not scale invariant. H. Sakoe, S. Chiba in [3] proposed a dynamic programming algorithm for speech recognition. These algorithms become increasingly popular because of the dynamic programming strategy can be combined with a very efficient and practical pruning strategy so that very large search spaces can be handled and secondly, it has turned out to be extremely flexible in adapting to new requirements with some limitations.

H. Ling, W.W. Jacobs in [4] proposed a method where the inner-distance is articulation insensitive and is good for complicated shapes with part structures. Then the inner-distance is used to build better shape representations but to compute the inner-distance the shape boundary is assumed to be known. This limits the approach to applications where the segmentation is available. R. da S. Torres, A.X. Falcão in [5] proposed two shape descriptors namely contour saliences and segment saliences for image retrieval and analysis. Loncaric in [6] dopts three different classifications proposed by avlidis in [7]. Shape matching methods can be either boundary or extitglobal, depending on whether they exploit only the silhouette or also the interior of the shapes. A number of shape matching techniques are based on some kind of shape skeletonization. Torres and Falcäo [5, 8] compute image skeletons at multiple scales and use them to detect salient points on the contour of the shape.

Instead of relying on shape skeletal points, some other global methods are based on the representation and the properties of all interior points of a certain shape. Gorelick [9] proposed the characterization of each interior point of the shape by the average distance that a random walker will travel before reaching it, assuming a starting point located on the shape's silhouette. Ebrahim [10] presented a method that transforms the raster of each shape to a one dimensional signal according to the occurrence of shape points on a Hilbert curve. This signal is then smoothed by keeping the largest coefficients of a wavelet transform.

The main issues regarding various problems of vision-based hand gesture recognition is to handle the large variety of gesture data. Recognizing gestures involve handling vast variances of the 2D appearance depending on the camera view point even for the same gesture, different silhouette scales and many resolutions for the temporal dimension. Moreover, it need also to go through various parameters such as the accuracy, performance, usefulness according to the type of application's use, robustness, scalability and user-independence.

III. PROPOSED APPROACH

In our proposed work, we implement image feature extraction by using LTrP texture pattern to extract texture features which gives histogram values for each pattern these gives thirteen different patterns. LTrP descriptor binary encodes the relation between the center pixel and its neighbors characterized by transformation consistency statistics of directional derivative in horizontal and vertical direction. After which binary encodes the magnitudes for the center pixel. Then it converts both binary code to decimal and computes the histogram traversing all pixels. Basically, LTrP descriptor is made up of two sections, that is, tetra pattern and magnitude pattern. Tetra pattern is captured based on first-order derivatives and transformation consistency statistics of directional derivative. Given one image, the first-order derivatives along horizontal and vertical direction at center pixel g_c can be evaluated using equation (1) and equation (2).

$$I^1_{hori}(g_c) = I(g_{k\_hori}) - I(g_c) \tag{1}$$

$$I^1_{verti}(g_c) = I(g_{k\_verti}) - I(g_c) \tag{2}$$

And the possible direction of each center pixel can be converted into equation (3).

$$I^1_{Dir(g_c)} = \begin{cases} 1, I^1_{hori(g_c)} \geq 0, I^1_{verti(g_c)} \geq 0 \\ 2, I^1_{hori(g_c)} < 0, I^1_{verti(g_c)} \geq 0 \\ 3, I^1_{hori(g_c)} < 0, I^1_{verti(g_c)} < 0 \\ 4, I^1_{hori(g_c)} \geq 0, I^1_{verti(g_c)} < 0 \end{cases} \tag{3}$$

Then tetra pattern for every pixel can be captured as using equation (4) and equation (5).

$$\begin{aligned} LTrP^2(g_c) = \{ & f(I^1_{Dir}(g_c), I^1_{Dir}(g_0)), \\ & f(I^1_{Dir}(g_c), I^1_{Dir}(g_1)), \\ & ...., f(I^1_{Dir}(g_c), I^1_{Dir}(g_{N-1})) \ \} \end{aligned} \tag{4}$$

$$f(I^1_{Dir}(g_c), I^1_{Dir}(g_k)) = \begin{cases} 0, & I^1_{Dir}(g_c) = I^1_{Dir}(g_k) \\ I^1_{Dir}(g_k) & otherwise, \end{cases} \tag{5}$$

Then SIFT algorithm performs matching of feature points and finds relevancy of it which gives result for classification. This feature concept is scale and rotational invariant feature vector and classifies the image accordingly.

SIFT feature concept is a type of feature used for classification in computer vision. SIFT feature point is the particular case of the texture model. It has been found to be a powerful feature for texture classification. The advantages of this approach are its performance related things and computational easiness, due to which it makes possible to analyse images in challenging real-time environments. Here feature points are detected by searching over all scales and image locations. Finding locations which are invariant to the scale change of the image can be accomplished by searching for stable features across all the scales, by using a function of scale which is known as scale space which consists of a group of blurred and subsampled versions of the original image. For an image, the scale space is constructed using a Gaussian kernel G(x, y, σ) with various values of σ.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{6}$$

where ∗ is the convolution operation in x and y and

$$G(x, y, z) = \frac{1}{2\Pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{7}$$

Starting with the original image , an initial step is to convolve the image with G(x, y, σ), which leads to a blurred image L(x, y, σ). This function is repeated using G(x, y, kσ) and it gives a further blurred image L(x, y, kσ). The difference between the two nearby blurred images is defined as a difference-of-gaussian image D(x, y, σ) which is given by

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{8}$$

The Fig. 2 shows the sequence of steps to perform the matching task. The flow of work according to the system architecture is first of all taking an input image from appropriate location, performing the preprocessing on it to reduce the noise from original image and to transform it into the appropriate size and resolution format, then doing the segmentation task to locate objects and boundaries in an image, after segmentation extracting the texture features from it by using LTrP technique, after that performing matching task using SIFT algorithm and finally recognizing the image and classifying it to the appropriate class.

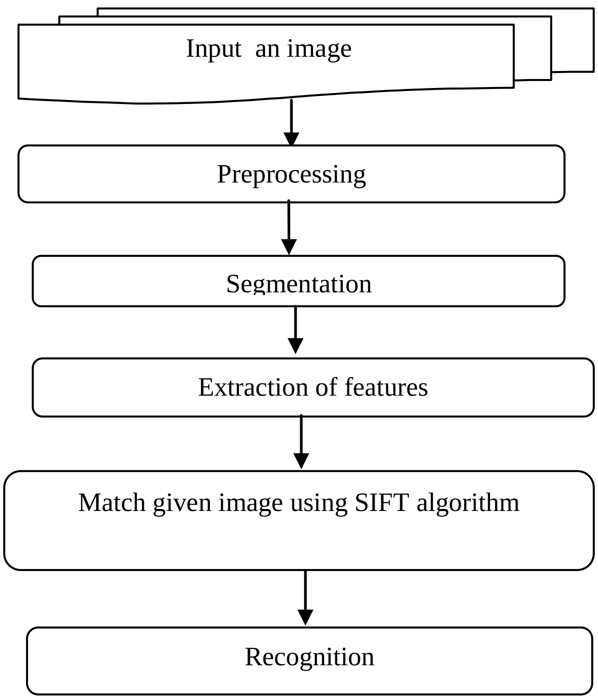

Fig. 2 System Architecture

IV. EXPERIMENTAL WORK AND RESULTS

Table I shows the comparative analysis of the result of proposed approach with some already developed techniques. For the purpose of experimental setup we have created our own database and trained it. To evaluate the results for proposed approach the evaluation parameters like sensitivity, specificity and accuracy has been calculated. These parameters are determined from the values of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). The formulae for accuracy, sensitivity and specificity are as follows

Accuracy = ((TP+TN) / (TP+TN+FP+FN))*100

Sensitivity = TP / (TP+FN) *100

Specificity = TN / (TN+FP) *100

TABLE I
PERFORMANCE EVALUATION METRICS

Approach	Method	Sensitivity	Specificity	Accuracy
Dr.E.Annasaro *et al.*in [11]	Canny	90.9337	80.7851	86.9174
	Sobel	87.4499	76.7933	83.3197
	Prewitt	86.9029	64.9826	76.6149
	Robert	82.1839	63.3776	74.08013
Damien Michel *et al.* in [1]	DTW	--	--	89.9000
Our Approach	LTrP + SIFT	92.3077	100	95.0000

Fig. 3 shows the comparative analysis of all the methods specified in table 1 in graphical notation.

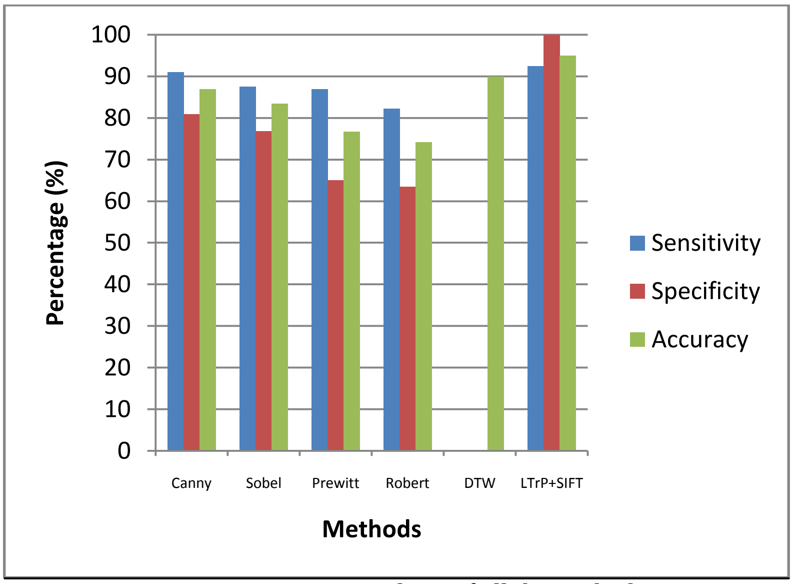

Fig. 3 Comparative analysis of all the methods

V. CONCLUSION AND FUTURE SCOPE

e propose a method for feature extraction to match and recognize the gesture of given image. We extract the texture features for the input image. The gesture region of a given template is defined as the set of all points that lie closer to this template than to any other of the available templates. The recognition method therefore consists in finding the template that is closest to the input vector. By using SIFT algorithm, we extract matching point of given image with database. The proposed approach provides a fine solution to the problem of partial shape matching. The key idea and main contribution of this paper lies in by using proposed technique increasing the accuracy and performance of the system. According to the results evaluated this technique has given the sensitivity 92.31%, specificity 100% and accuracy up to 95%. Finally, these comparative experimental results demonstrate the effectiveness and usefulness of the proposed method compared to existing ones.

This work may extend to make actions according to eye movement, secondly we can increase number of features to be matched, perform matching over 3D images also.

ACKNOWLEDGMENT

Authors would like to thank UGC for giving the facility UGC SAP (II) DRS Phase-I F.No.3-42/2009 in the Department of Computer Science and Information Technology and partially support to this work.

REFERENCES

[1] Damien Michel, Iasonas Oikonomidis and Antonis Argyros, "Scale invariant and deformation tolerant partial shape matching", Image and Vision Computing 29 (2011) 459–469.

[2] T.B. Sebastian, P.N. Klein and B.B. Kimia, "Recognition of shapes by editing shock graphs", IEEE International Conference on Computer Vision, 2001, pp. 755–762.

[3] H. Sakoe and S. Chiba, "A dynamic programming approach to continuous speech recognition", Proceedings of the 7th Int'l Congress on Acoustics, Budapest, 1971.

[4] H. Ling and W.W. Jacobs, "Shape classification using the inner-distance", IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (2) (Feb. 2007) 286–299.
[5] R. da S. Torres and A.X. Falcão, "Contour salience descriptors for effective image retrieval and analysis", Image and Vision Computing Journal 25 (1) (2007) 3–13.
[6] S. Loncaric, "A survey of shape analysis techniques", Pattern Recognition 31 (1998) 983–1001.
[7] T. Pavlidis, "A review of algorithms for shape analysis", Computer Graphics and Image Processing 7 (2) (1978) 243–258.
[8] R. da S. Torres, A.X. Falcão, L. da Fa Costa, "A graph-based approach for multiscale shape analysis", Pattern Recognition 37 (2004) 1163–1174.
[9] L. Gorelick, M. Galun, E. Sharon, R. Basri and A. Brandt, "Shape representation and classification using the poisson equation", IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (12) (2006) 1991–2005.
[10] Y. Ebrahim, M. Ahmed, W. Abdelsalam and S.C. Chau, "Shape representation and description using the Hilbert curve", Pattern Recognition Letters 30 (4) (2009) 348–358.
[11] Dr.E. Annasaro and A. Hema, "Color and shape feature extraction and matching in pill identification systems", International Journal of Computer Science and Information Technologies, Vol. 5 (2), 2014, 1011-1015.

