Asian Journal of Applied Science and Technology (AJAST)
Volume 1, Issue 2, Pages 268-273, March 2017

Object Recognition Using Shape Context with Canberra Distance

K.Senthil Kumar

Assistant Professor, Department of Electrical and Electronics Engineering, Tamilnadu College of Engineering, Coimbatore, India.

Article Received: 13 March 2017 *Article Accepted: 23 March 2017* *Article Published: 26 March 2017*

ABSTRACT

Object recognition is the challenging problem in the real world application. Object recognition can be achieved through the shape matching. Shape matching is preceded by i) detecting the edges of the objects from the images. ii) Finding the correspondence between the shapes. iii) Measuring the dissimilarity between the shapes using the correspondence. iv) Classifying the object into classes by using this dissimilarity measures. In order to solve the correspondence problem, we attach a descriptor, the shape context, to each point. The shape context at a reference point captures the distribution of the remaining points relative to it, thus offering a globally discriminative characterization. Corresponding points on two similar shapes will have similar shape contexts. The one to one correspondence is achieved through the cost based bipartite graph matching. The cost matrix is reduced through Hungarian algorithm. The dissimilarity between the two shapes is computed using Canberra distance. The nearest neighbor classifier is used to classify the objects with the matching error. The results are obtained using the MATLAB for MINIST hand written digits.

Keywords: Correspondence, Digit recognition, MNIST Database and Shape Context.

1. Introduction

Object Recognition in real time systems is a complex problem and which is one of the common problem featured in computer vision applications. Shape matching is an important ingredient in shape retrieval, recognition and classification. The shape matching begins with the image processing. The steps involved in the shape matching such as image acquisition, preprocessing are done with the help of image processing techniques. The morphological operation in image processing plays a vital role in the shape matching and object recognition. The morphological operation is a tool for extracting image components that are useful in the representation and description of region.

The process of identifying objects in images is called Pattern Recognition (PR). According to Morton [15] Pattern is a distinctive arrangement of structural elements. Pattern means archetype or prototype according to which objects are formed. Identification is the act of assigning an object to the class patterns to which it belongs. It can be called as Pattern identification. It is one of the key techniques used in Artificial Intelligence (AI), which is concerned with problem solving.

The pattern recognition and the image processing are interrelated. The main application areas are Optical Character Recognition (OCR), Medical imagery and medical diagnosis, automation of aerial and satellite photo interpretation, industrial automation and inspection and model based recognition to detect defects in integrated circuits.

The ultimate goal of computer vision developed by David [7] is to use computers to emulate human vision and take actions based on visual inputs. This area itself is a branch of AI whose goal is to emulate human intelligence. A Machine Vision (MV) system proposed by Ramesh [18] is to create a model of the real world from the images. It recovers useful information about a scene from its two dimensions. Since the images are two dimensional projections of three dimensions, the information is not directly available and must be recovered. The Machine Vision algorithms take images as inputs but produce other types of outputs. The information are recovered automatically and not as image processing.

The structure of this paper is as follows: we discuss the related work in Section 2. In Section 3 we describe the proposed method in detail. Section 4 gives the results and data base for hand written digits. We conclude in Section 5.

2. METHODOLOGY

The shape can be defined as the equivalence class under a group of transformations. The shape matching is identifying the shape when two shapes are exactly the same. The statistician's definition of shapes is addresses the problem of shape distance, assumes the correspondence are known. The shape matching can identify without finding the correspondence by using the intensity based technique. An extensive survey of shape matching in a computer can broadly classified into two approaches.

1. Brightness based method
2. Feature based method

2.1 Brightness Based Method

Brightness based or appearance based method is a complementary view to feature based methods. Instead of focusing on the shape this approach make direct use of gray values within the visible portion of the objects. The brightness information is used to find the correspondences and to align the gray scale values to compare brightness of two different shapes. The Fitting Hand Craft Model (FHC) presented by Yuille [22] suggest a flexible model to build invariance to certain kinds of transformations, but it suffers from the need of human designed templates and the sensitivity to initialization when searching via gradient descent.

Elastic graph matching developed by Lades et al [16] which involves both the geometry and photometric features in the form of local descriptors based on the Dynamic Link Architecture, which is a Gaussian derivative. The performance is evaluated using statistical analysis. The alternative is the feature based method.

The way by building classifiers to find correspondence have learning algorithm which produce good results by using Principal Component Analysis (PCA) when used in probabilistic methods for face recognition.

2.2 Feature Based Method

Feature based method is done by using the boundaries of silhouette images. It uses only the outer markings. Boundaries are conveniently represented by arc length. The dynamic programming approach is used, which is fast but the drawback is that it does not return the correspondences. The advantage of the algorithm is fast and invariant to several kinds of transformation including some articulation and occlusion. It uses edge detector. It is only for 2D image.

There are several approaches for shape recognition based spatial configurations of small number of key points. These models are used to vote for a model without solving the correspondence. The following techniques are based on the feature based method.

On growth and form model developed by Thompson [21] observed that the related or similar but not identical shapes can often be formed into alignment using simple coordinate transformation. This work illustrates that the related shapes has the unique structure but it differs on the properties like scaling, rotation, shearing and projections. So it can't be used directly by applying the corresponding transformation.

Mass spring model proposed by Fischler and Elschlager [6] reviewed the on growth and form model, and operationalized such an idea by means of energy minimization in a mass spring model. The mass spring model minimizes the energy by using the dynamic programming technique. The new programming system does not need to be written for every new description, one just specifies description in terms of certain set of parameter.

Thin Plate Spline (TPS) is an effective tool modeled by Bookstein [4] for modeling coordinate transformation in several computer vision applications. In practical applications, a good solution is possible only by using a small subset of corresponding points. By using the approach of subset, the transformations for all the points be estimated. The technique from the machine learning for function approximation using Radial Basis Functions (RBF) is adapted to the task, which shows a significant improvement over the naive method. One drawback of this method is that it does not allow for principal warp analysis. By means of experiments on real and synthetic data, the pros and cons of different approximations are demonstrated so as to take decision suited for application.

Fitting hand craft model designed by Yuille [24] is an approach for extracting facial features from images and also for determining the spatial organization between the features using the concept of a deformable template. This is a parameterized geometric model of the object. Variation in the parameters corresponds to deformation of the object are specified by a probabilistic model. After the extraction stage the parameters of the deformable template can be used for object description and recognition.

Active Shape Model (ASM) is a robust approach to recognize and locate known rigid objects in the presence of noise, clutter, and occlusion. It is more problematic to apply model based method to images. The problem with existing methods is that they sacrifice model in order to accommodate variability, there by compromising robustness during image interpretation. Cootes et al [5] describes a method for building models by learning patterns of variability from a training set of correctly annotated images, which can be used for image search in an iterative refinement algorithm that employed by Active Shape Models. The result shows that the method for constructing models combined with the active matching technique provides a systematic and effective step for the interpretation of complex images.

Shape Features and Tree Classifiers introduced by Amit et al [1] describe a very large family of binary features for two dimensional shapes. The feature of separating particular shape family is determined from the training samples during the construction of classification trees. Different trees correspond to different aspect of shapes. Adopting the algorithm to a shape family is fully automatic, once the trained samples are provided. The best error rate for a single tree is 0.7 percent and the classification rate is 98 percent. The standard method for constructing tree is not practical because the feature set is virtually infinite.

Shape context presented by Belongie [3] is a novel approach in which the similarity between shapes is measured and used it for object recognition. The measurement of similarity is done by solving for correspondences between the shapes. The shape context at a reference point captures the remaining points relative to it, thus offering a character. Corresponding points on two similar shapes will have similar contexts. By using the transformation, matching between the two shapes is performed. The dissimilarity between the two shapes is given by sum of matching errors between corresponding points and the results are presented for silhouettes, handwritten and coil data sets.

Robust shape similarity retrieval developed by Attalla and Siy [2] is a novel shape retrieval algorithm that can be used to match and recognize 2D objects. The matching process uses a new multi-resolution polygonal shape descriptor that is invariant to scaling, rotation and translation. The shape descriptor equally segments the contour of any shape, regardless of its complexity and captures three features around its center including the distance and slope relative to the center. The novel shape matching algorithm uses the shape descriptor and applies it by linearly scanning to a stored set of shapes and measures the similarity using elastic comparisons. The multi-resolution segmentation provides flexibility for applications with high accuracy results and the elastic matching adds an advantage when matching for partially occluded shapes. The algorithms are tested on many databases.

Contour flexibility is a technique developed by Jianzhuang Liu [9] for shape matching in which the predominant problem is matching the shapes, when articulation and deformation occurs. These variations are insignificant for human recognition, but the matching algorithm give results that are inconsistent with our perception. A novel shape descriptor of planar contours, called contour flexibility, which represents the local and global features from the contours. The shape of an object is important in object recognition. A small change in size of the object results insignificant changes. The valuable features for shape recognition mostly come from the corners or the part where the changes occur.

The drawbacks of the existing methods are it does not return correspondences, suffers from the need for human designed templates and applied for 2D images and limited 3D images.

3. PROPOSED METHOD

The object can be treated as point set. The shape of an object is essentially captured by a finite subset of its points. A shape is represented by a discrete set of points sampled from the internal or external contours on the object.

These can be obtained as location of pixels as found by an edge detection. These shapes should be matched with similar shapes from the reference shapes. Matching with shapes is used to find the best matching point on the test image from the reference image.

The proposed algorithm for matching with the shapes contains the following steps:

1. Preprocessing
2. Feature extraction
3. Finding correspondence
4. Applying transformation
5. Similarity measures
6. Classifying images

3.1 Finding Correspondence

Finding correspondence is measuring the dissimilarity between the reference image and the test image. In finding correspondence the statistical method is applicable. Statistical method is used to find the distance between the reference image and the test image.

Shape Context

For each point p_i on the first shape, find the best matching point q_j on the second shape. This is a correspondence problem similar to that in stereopsis. Consider the set of vectors originating from a point to all other sample points on a shape. These vector express the configuration of the entire shape relative to the reference point. The shape is represented by the n-1 vector, since n gets larger the shape become the exact. By this context, the distribution of pixel comes to known.

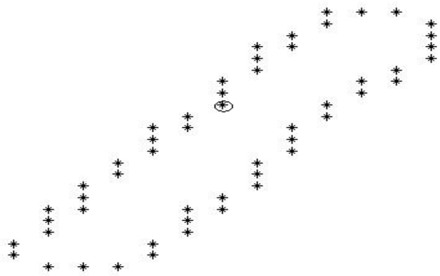


Fig 3.1 Sampled edge of one

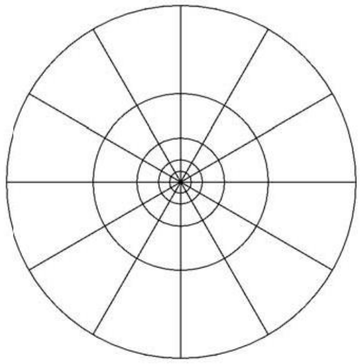


Fig 3.2 log polar histogram bin for computing shape context

For a point p_i on the shape, we compute the course of histogram h_i of the relative coordinates of the remaining n-1 points. The figure 3.1 shows the shape of the digit one.

The bins that is uniform in log-polar space, making the descriptor more sensitive to positions of nearby sample points than to those of points further away. The figure 3.2 shows the structure of the bin in which 5 radius and 12 angles are used.

The cost matrix formed for a point p_i and q_j on the second shape by using the chi square test. The chi square test for the point p_i and q_j is given by,

$$c_{ij} = \frac{1}{2}\sum_{i=1}^{n-1}\frac{(h_i(k) - h_j(k))^2}{(h_i(k) + h_j(k))}$$

Where $h_i(k)$ and $h_j(k)$ denote the K-bin normalized histogram at p_i and q_j respectively. The cost matrix is reduced by Hungarian algorithm.

3.2 Transformations

To reduce the error, the transformations are applied to the image. The transformation can defined as the changing the image from one form to another. An affine transformation is a transformation that preserves collinearity (i.e., all points lying on a line initially still lie on a line after transformation) and ratios of distances (i.e., the midpoint of a line segment remains the midpoint after transformation). The transformation can perform shearing, scaling, rotation, translation and shifting. Hence the affine transformation can match the images which are changing at these parameters (rotation, scaling, shearing). The transformation transform the image into different shifting therefore this transformed image will reduce the error considerably.

3.3 Dissimilarity Measures

The dissimilarity measured through the Canberra distance. After finding the corresponding point between the shapes, the dissimilarity of the image find out through the Canberra distance as in equation below.

$$D(x,y) = \sum_{i=1}^{n} \frac{x_i - y_i}{|x_i| + |y_i|}$$

Here the value of n=2since it is a two dimensional and xi represents point on first shape y_i represents corresponding points on the second shape.

3.4 Classifying Images

The images are classified into the different classes by using the matching error. The k-NN classifier is used to classify the images. In this classifier, the matching error is given with 10 centered classes, depending on the closeness of the matching error with the centered value the image has classified.

4. RESULTS

The algorithm tested with the MNIST database, and developed with the help of the MATLAB. MNIST database serves for comparison of different methods of handwritten digit recognition. So the MNIST database is used for testing and training the algorithm. The MNIST database contains 60,000 handwritten digits in the training set and 10,000 handwritten digits in the test set. The MNIST database is shown in figure 4.5.

The time required is also a parameter to identify the algorithm efficiency. It can be measure by calculating the time required for identification of the single character.

Time Required=1.6s.

The following figures show that the input and processed output images.

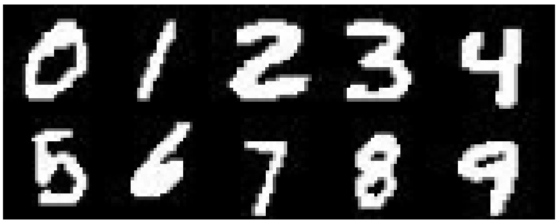


Figure 4.1 Input Images

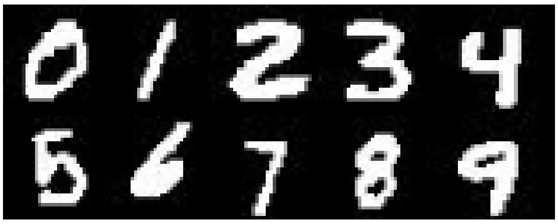


Figure 4.2 Noise Cancelled Images

Figure 4.3 Edge Detected Images

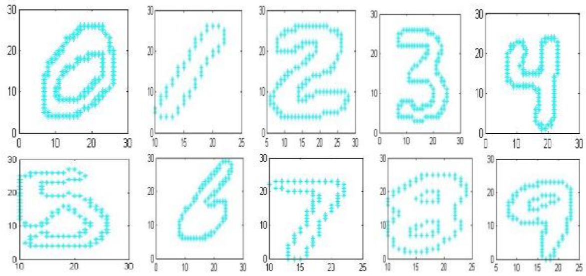
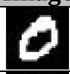
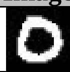
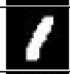
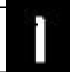
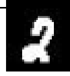
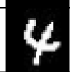
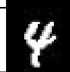
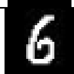
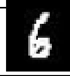
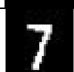
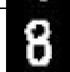


Figure 4.4 Mapped Images

TABLE I: Comparison of the matching error with the Euclidian distance and Canberra distance.

Input Image	Reference Image	Hungarian method	Modified Hungarian
		2.367	2.243
		2.123	2.002
		2.231	2.201
		2.434	2.241
		2.125	2.012
		2.211	2.391
		2.764	2.543
		2.291	2.151
		2.312	2.111
		2.956	2.765

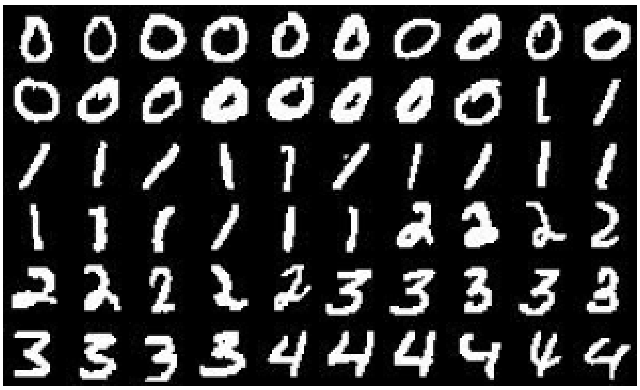


Figure 4.5 MNIST database

5. CONCLUSION

This paper gives the efficient method for finding the class of the hand written images. This method is simple and the invariant to the noise and the error is also reduced. This can also be extended to the hand written characters, industrial objects, face recognition and pedestrian identification. Further, the algorithm will extend to recognize multiple objects from the image simultaneously. It can improve the reliability of the real time system. The algorithm can also be applied for the video image and aerial images to identify the object. The proposed method can be developed for different applications like image retrieval, military areas, investigation departments, and industrial automation.

REFERENCES

[1] Amit Y., Geman D. and Wilder K., 'Joint Induction of Shape Features and Tree Classifiers' *IEEE Transaction on Pattern Analysis and Machine Intelligence,* Vol 19, pp. 1300 – 1305, 1997.

[2] Attalla E. and Siy P., 'Robust Shape Similarity retrieval Based on Contour Segmentation Polygonal Multi resolution and Elastic Matching' *Pattern Recognition,* Vol 38 (12), pp. 2229 – 2241, 2005.

[3] Belongie S., Malik J. and Puzicha J., 'Shape Context: A new descriptor for shaping and object recognition' *IEEE Transaction on Pattern Analysis and Machine Intelligence,* Vol 24, pp. 831 – 837, 2002.

[4] Bookstein F.L., 'Principal Warps: Thin-Plate Splines and Decomposition of Deformations' *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol 11, no 6, pp. 567 – 585, 1989.

[5] Cootes T., Cooper D., Taylor C. and Graham J., Active Shape Model - their training and Application' *Computer vision and image understanding,* Vol 61, pp. 38 – 59, 1995.

[6] Fischler M. and Elschlager R., 'The Representation and Matching of pictorial structures' *IEEE Transaction Computers,* Vol 22, no 1, pp. 67 – 92,1973.

[7] David Forsyth A., 'Computer Vision' *First Indian Edition, Pearson Education,* 2003.

[8] George W., 'Statistical Methods' *sixth Edition, Oxford IBH publication,* 1967.

[9] Jianzhuang Liu, '2D Shape Matching by Contour Flexibility' *IEEE Transaction on Pattern Analysis and Machine Intelligence,* Vol 19, pp. 1 – 7, 2008.

[10] Klassen E., Srivastava A., Mio W. and Joshi S.H., 'Analysis of planar shapes using geodesic paths on shape spaces' *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol 26 (3), pp. 372 – 383, 2004.

[11] Kunttu I., Lepisto L., Rauhamaa J. and Visa A., 'Multistate Fourier Descriptor for Shape Classification' *Proc. Int'l Conf. on Image Analysis and Processing,* Vol 1, pp. 536 – 541, 2003.

[12] Ling K. and Jacobs D.W., 'Using the Inner-Distance for Classification of Articulated Shapes' *IEEE Conference on Computer Vision and Pattern Recognition,* Vol 2, pp. 719 – 726, 2005.

[13] Lowe D.G., 'Object Recognition from Local Scale - Invariant Features' *Proc. Seventh International Conference on Computer Vision,* pp. 1150 – 1157, 1999.

[14] Milios E. and Petrakis E., 'Shape Retrieval Based on Dynamic Programming' *IEEE Transactions on Image Processing,* Vol 9 (1), pp. 141 – 147, 2002.

[15] Morton Nadler, 'Pattern Recognition Engineering', *John Wiley & Sons,* 2000.

[16] Lades M, Vorbuggen C. and Lange J., 'Distortion Invariant Object Recognition in the Dynamic Link Architecture' *IEEE Transactions on Computers,* Vol 42, no 3, pp. 300 – 311, 1993.

[17] Rafael Gonzalez C., 'Digital Image Processing' *Second edition, Prentice Hall Publication,* 2006.

[18] Ramesh Jain, Rangachar Kasturi and Brian Schunck G., 'Machine Vision', McGraw-Hill International Editions, 1995.

[19] Shpiro L.S. and Brady J.M., 'Feature-based correspondence: An Eigenvector Approach' *Image and Vision Computing,* Vol 10, no 5, pp. 283 – 288, 1995.

[20] Super B.J., 'Learning Chance Probability Functions for Shape Retrieval or Classification', *IEEE Workshop on Learning in Computer Vision and Pattern Recognition,* Vol 6, pp. 93 – 98, 2004.

[21] Thomson D.W., 'On Growth and Form', *Cambridge Univ. press,* 1917.

Volume 1, Issue 2, Pages 268-273, March 2017

[22] Tu Z. and Yuille A.L., 'Shape Matching and Recognition-Using Generative Models and Informative Features', *Proc. European Conf. on Computer Vision*, pp. 195 – 209, 2004.

[23] Vetter T., Jones M.J. and Poggi T., 'A Bootstrapping Algorithm for learning linear models of object classes' *IEEE Conference on Computer vision and Pattern Recognition,* pp. 40 - 46,1997.

[24] Yuille A., 'Deformable Templates for Face Recognition' *Journal of Cognitive Neuroscience*, Vol 3, no 1, pp. 59 – 71, 1991.

[25] Zahn C. and Roskies R., 'Fourier Descriptors for Plane Closed Curves' *IEEE Transaction Computers,* Vol 21, no 3, pp. 269 – 281, 1972.