International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 04 | Apr 2020 www.irjet.net p-ISSN: 2395-0072

Object Detection using Hausdorff distance

Jayesh Bhadane[1], Chetan Mhatre[2], Arvind Chaudhari[3]

[1,2,3]*Department of Computer Engineering, Pillai HOC College of Engineering and Technology, Rasayani, India*
[4]*Prof. Jacob John, Department of Computer Engineering and Technology, Pillai HOC College of Engineering and Technology, Rasayani, India*

Abstract - *The need for a proper and accurate system of object recognition is increasing in our day to day life. Such systems can be used in developing driver-less cars, extracting information from handwritten documents, reading characters and digits, etc. The famous YOLO algorithm works well for object recognition with a speed of 155 frames per second but is less accurate. The YOLO model struggles with small objects that appear in groups, such as a flock of bees. Also, another model uses Euclidean distance for recognizing objects. But for some of the objects, it doesn't work well. There are situations where the model captures the garbage data in random noisy situations. In the proposed work, a new object recognizing system is developed with the help of Hausdorff distance. The new Hausdorff distance can dispose of the noise desirably. Encouraging results were obtained after testing the algorithm with MNIST, and the COIL dataset along with a private collection of handwritten digits.*

***Key Words*: Shape matching, Hausdorff distance, Digit recognition, MNIST database.**

1. INTRODUCTION

The difficulty of object recognition has been researched for more than forty years and impressive consequences have been obtained. Unfortunately, the hand written digit reputation nonetheless the real demanding situations to researchers. The hand written digit is semi-cursive, where every digit can be written in many of form. Figure 1 shows the identical digit five in amazing forms. Shape evaluation is a fundamental problem in computer imaginative and prescient and image processing, and affects a whole lot of application domains. It is crucial key issue to apprehend the gadgets in the image.

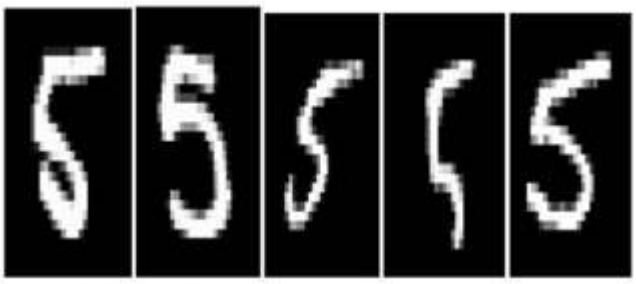

Fig -1: Example of hand written digit.

The geometry houses of the picture provides the robust signature for recognition below the noisy environment, and additionally the shape of the item has been used for matching the input picture to the templates. Shape can be represented by a set of points. It is an efficient method because of its low dimensionality [1]. Alternatively, the shape can also be represented as the regions enclosed by the planar curves [1][4]. It may be efficient underneath topological changes, and invariant to noise. The efficient characteristic need to be invariant beneath transformation. Since the shape is a property of geometric gadgets, it is invariant below transformation. The geometry homes of shape can be represented by using a statistical feature seeing that it offers with the point set. It leads to the efficient evaluation of geometric gadgets. Belongieet al proposed a novel shape descriptor, called shape context which represents the shape explicitly with a radial histogram [2]. The set of rules checked with MNIST, coil database and MPEG records bases and shows the better recognition rate. Attneave represents the shape by curvatures as a shape descriptor due to its invariant houses and computational convenience [8]. The unwanted parameterization of shape matching is avoided with the aid of the use of signature that consists of curvature and its first derivative. Pedro F. Felzenszwalb defined an object detection machine based on combos of multi scale deformable element models [5]. The system relies upon on discriminative training of classifiers. It is one of the efficient techniques for matching deformable models to image. Proposed work recognises object employing a form context and Hausdorff distance is introduced. In the start, the form context is computed for 2-point set and Hungarian algorithm is employed to search out the correspondence between 2 them. The process evaluates the similarity of the two-point set using Hausdorff distance [3]. The method was tested with the MNIST, COIL knowledge sets and a personal assortment of hand written digits and inspiring results were obtained.

2. RELATED WORKS

There have been many works on object detection both with and without Hausdorff distance. YOLO approach is developed for object detection that binds the objects with a box and labels the name of the object that it has identified. It is stated that it is more accurate when compared to traditional systems. Text recognition is finding a new pavement in machine learning techniques. Gomez et al. have proposed a selective search algorithm for identifying the texts in a word document. The algorithm generates a hierarchy of word hypothesis and produces an excellent recall rate when compared to other algorithms. Experiments conducted on ICDAR benchmarks demonstrated that the novel method proposed in extracted the textual scenes from

the natural scenes in a more efficient way. When compared to conventional approaches, the proposed algorithm showed stronger adaptability to texts in challenging scenarios. Convolutional Neural Networks used in Biomedical Image Segmentation enable precise localization of neuronal structures when observed in an electron microscopic stacks. The selective search algorithm for detecting the objects was used to generate possible object locations. The Selective Search algorithm shows results in a small set of data-driven, class-independent, high-quality locations. Gupta A, generated a series of synthetic data that was scalable and very fast. These images were fully used to train a Fully Convolutional Regression Network that efficiently performed text detection and bounding box regression at all locations and multiple scales in an image. The end model was able to detect the texts in the network significantly out claimed that it outperforms current methods for text detection in natural images.

TextBoxes++ as a single shot Oriented Scene Text Detector which detects arbitrary oriented scene text with both high accuracy and efficiency in a single network forward pass. There was no post-processing process and also had an efficient non maximum suppression. The proposed model was evaluated on four public data sets. In all experiments, TextBoxes++ claimed that it outperformed competing methods in terms of text localization, accuracy, and runtime. A novel approach was proposed in named Cascaded Localization Network (CLN) that joined two customized convolution nets and used it to detect the guide panels and the scene text from a coarse to fine manner. The network had a popular character-wise text detection and was replaced with string-wise text region detection, that avoided numerous bottom up processing steps such as character clustering and segmentation. Instead of using the unsuited symbol-based traffic sign datasets, a challenging Traffic Guide Panel dataset was collected to train and evaluate the proposed framework. Experimental results demonstrated that the proposed framework outperformed multiple recently published text spotting frameworks in real highway scenarios. Guo et al. proposed a cost-optimized approach for text line detection that worked well for documents that were captured with flat-bed and sheet-fed scanners, mobile phone cameras, and with other general imaging assets. Sliding-window based mostly ways are terribly popular within the field of scene text detection. Such ways create use of the native structure property of text and scan all potential positions and scales within the image. The algorithmic program projected during this paper conjointly works during a sliding window fashion. The main difference is that previous ways ask for scene text either at a reasonably coarse graininess or at a fine graininess, whereas our algorithmic program capture scene text at a moderate graininess (several adjacent characters. The advantages are:

1) It permits to exploits the symmetry property of character teams, that cannot be excavated at stroke level.
2) It will handle variations inside a word or text line, like mixed case and minor bending.

3. PROPOSED METHODOLOGY

3.1 Pre-processing

The pre-processing approach is editing the image for quality matching to the reference image. Noise cancellation is one of the pre-processing techniques. For noise cancellation a number of methods can be used like Median filter, Mean filter, Gaussian clear out. The Gaussian filter out has been used for noise cancellation. Gaussian filters are a class of linear smoothing filters with the load chosen in keeping with the form of the Gaussian function. The Fig. 2 indicates the threshold detected image of digit 2. The Gaussian smoothing filter out is excellent filtering for eliminating noise drawn from a ordinary distribution. The object can be handled as point set. The shape of an object is largely captured by a finite subset of factors. A shape is represented with the aid of a discrete set of points sampled from the inner or external contours of the object. Contours may be received as area of pixels as located through an edge detection. The shapes should be matched with similar shapes from the reference shapes [6]. Matching with shapes is used to locate the exceptional matching factor on the check image from the reference image.

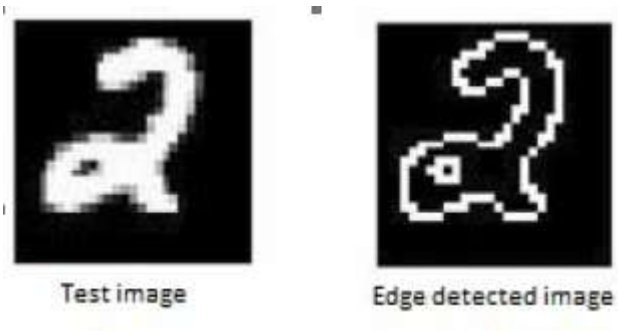

Fig -1: Test image and reference image

3.2 Feature Extraction

The image needs to be compared for the unique capabilities like wide variety of pixel, width, length, edges and brightness of the picture. The imaginative and prescient processing identifies the features in image that are applicable to estimate the structure and houses of objects in a photograph. Edges are one such feature. The edges can be used as the unique functions of the input and the reference photo for its simplicity and effective matching. The canny aspect detector is used for the detection of the rims of the photo. The canny edge detector is the first derivative of a Gaussian and carefully approximates the operator that optimizes outcome of signal to noise ratio and localization [7]. Detection of side by way of lesser error rate, which means with the purpose of the detection must efficiently catch as numerous edges shown inside the photo as likely. The area factor detected from the operator must correctly localize on the center of the area. A given aspect inside the photograph must only be marked once, and where potential, picture noise must no longer generate false edges.

3.3 Correspondence

Correspondence is observed for the two-factor set by the usage of the form descriptor, shape context and shortest path augmenting algorithm. Further, with this correspondence, the aligning rework is implemented at the second shape with reference to the first one. The new release continues till the quality matching occurs or maximum generation reaches. In this algorithm, the maximum 5 iterations are processed for the best matching.

3.4 Hausdorff Distance

Hausdorff distance from set A to set B is a maximin function, defined as the directed Hausdorff distance function h(A,B) = max(min(D(A,B)) Where a and b are points of units A and B respectively, and d(a, b) is any metric between these factors. A more general definition of Hausdorff distance could be H (A,B) = maxh(A,B),h(B,A). The distances h(A, B) and h(B, A) are from time to time termed as ahead and backward Hausdorff distances of point set A to factor set B. The Hausdorff distance used to calculate the distance from the test image and the reference pictures inside the different class. The distance gives the similarity records of test shape to the form within the distinct classes. The fig. 3 below suggests the only to one matching between the test and reference image. It indicates the Hausdorff distance is 1.4371.

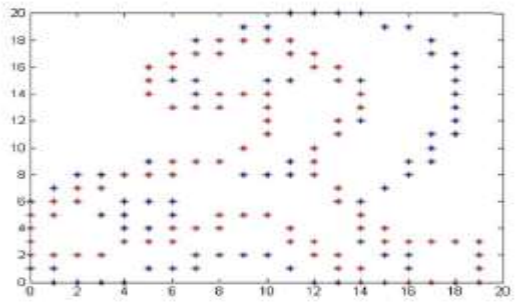

Fig -1: One to one matching between test image and reference image

3.5 Classification

The NN class is used to categorise the check digits into the extraordinary instructions with ten quantity of input layer, 15 variety of hidden layer and 10 variety of output layer. The precept at the back of nearest neighbor methods is to discover a predefined quantity of training samples closest in distance to the new point, and predict the label from these [3]. The wide variety of samples in proposed method is two for each digit. The Hausdorff distance is used as metric measures [3].

4. PROPOSED ALGORITHM

The following algorithm indicates the step by step process for the shape matching using Hausdorff distance.

1) Acquiring the images.

2) Removing the noises from image.

3) Finding the edges on the image and it can be treated as shape of the object.

4) Matching the test shape and reference shapes from the data base using the Hausdorff distance.

5) Based on the distance the shapes are classified into the object.

The following flowchart depicts the shape matching using Hausdorff distance.

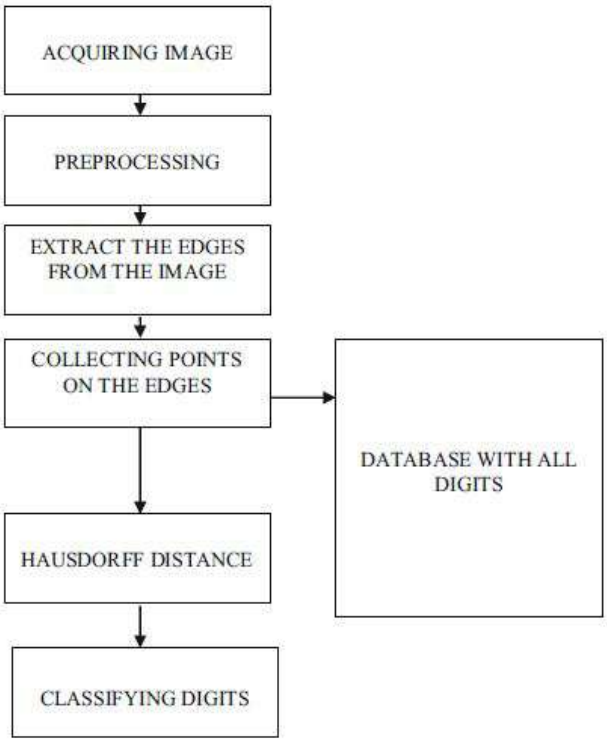

Fig -1: Flowchart of shape matching using Hausdorff distance.

5. RESULTS

The algorithmic program is checked with the MNIST and COIL information. The MNIST information consists of the 2-D written image. The sides and bounds are simply known, so 100 points from edges are taken for form matching. Whereas within the COIL information the 3D images were used thus, the interior and external edges were taken for form matching. Three hundred edge points were thought-about for the COIL information form matching.

5.1 MNIST dataset

The algorithmic rule is checked with the MNIST information. The MNIST information consists of 60,000 training image and 10,000 test images. The information within the single

record, the hand written pictures are hold on with the dimensions of 784 in single row that has the image size of 28 × 28. Fig. 5 shows some sample images of MNIST database. The algorithmic rule gives the cheap results compare to the prevailing strategies. The algorithmic rule acknowledges 9968 pictures properly out of 10,000 pictures from MNIST dataset. The error rate is reduced to 0.72 percent. The program is run with MATLAB 14b on 2 GB RAM machine. The time required for single match is 1.47 s. Syed Salman Ali et al. developed the system with DCT and HMM though the speed of the system increases the error rate reduced [8]. Since the proposed method yields better result compared to other methods, it may be extended for other dissimilarity measures.

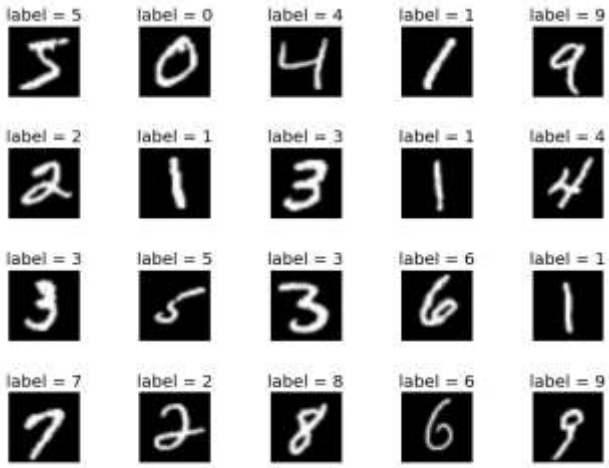

Fig -1: Sample prototype in MNIST database.

5.2. COIL dataset

The algorithm is applied for the COIL database. In COIL-20 database incorporates the photos of real-world gadgets inside the one of a kind of angle. The sampling purpose on the image is elevated from 100 to a few hundred points every inner and external contour are taken into the thought. The prototypes are included inside the algorithms. The prototype choice is a critical aspect that affects the recognition of the items. The training set is prepared through taking the number of equally spaced views for each object. The remaining items are taken because the testing snap shots. The prototype photos were taken because the reference image rather of single reference image. Fig. 6 shows the sample prototype images used COIL database. The input picture is as compared with all the reference photographs and the minimum blunders is taken for the class. The nearest neighbour classification is used for class. 47 images were misclassified out of 1200 pics. The errors charge is 3.91.

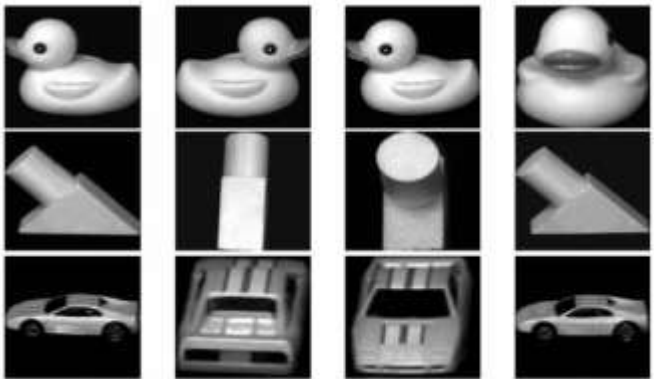

Fig -1: Sample prototype in COIL database.

6. CONCLUSION AND FUTURE WORK

This proposed method produces efficient results that too in reduced timing. Besides, the complexity of the program is also very less. This methodology can be extended to 3-d images and industrial image set. It can also be used be to recognize objects in videos. This methodology has multiple application in military areas, industrial applications, medical operations (when worked with an enhanced medical image dataset), investigation departments.

REFERENCES

[1] Forss´en PE, Lowe DG (2007) Shape descriptors for maximally stable extreme regions. In 2007 IEEE 11th International Conference on Computer Vision, p 1–8.

[2] Belongie S, Malik J, Puzicha J (2002) Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(4):509–522.

[3] K. Senthil Kumar, T. Maniganda, D. Chitra, L. Murali "Object recognition using Hausdorff distance for multimedia applications", Tamilnadu, College of Engineering, Coimbatore, India, June 2019.

[4] Prasad BG, Biswas KK, Gupta SK (2004) Region-based image retrieval using integrated color, shape, and location index. Comput Vis Image Underst 94(1–3):193–233

[5] Hong B-W, Soatto S (2014) Shape matching using multiscale integral invariants. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[6] Amit Y, Geman D,Wilder K (1997) Joint induction of shape features and tree classifiers.

[7] McIlhagga W (2011) The canny edge detector revisited. Int J Computer Vis 91.

[8] Ali SS, UsmanGhani M (2014) Handwritten digit recognition using DCT and HMMs. 12th International Conference on Frontiers of Information Technology, vol. 1.

[9] Attneave F (1954) Some informational aspects of visual perception. Psychol Rev 61(3):183