A STUDY OF DIFFERENT TECHNOLOGIES USED FOR CHARACTER AND FACE RECOGNITION

[1] Ritu Mathur, [2] ManjuMathur
[1] Student, M.tech, [2] Assistant Professor
[1,2] Digital Communication,
[1,2] Rajasthan College of Engineering for Women, Jaipur, India

Abstract : This paper is a review paper about various techniques used for character and face recognition applications. This paper describes about the performance of different technologies for Optical Character Recognition (OCR) and face recognition, in various conditions like low quality image, distorted image etc. We know that as the quality of the image varies, the performance of technologies changes. That's why we discuss all the possible distractions for various conditions.

IndexTerms **- Optical Character Recognition, Face Recognition, Templates, SVM.**

I. INTRODUCTION

As we all know that in today's world demands for new inventions and technologies have been arrived so rapidly that it becomes very important to transform old technologies into new versions. As we know that in today's world digital technologies are rapidly increasing. This increases need for digital technology for reading characters from images. As the demand for high security is increasing, Face Recognition is becoming an essential technique for authentication.

These types of technologies are widely used in commercial applications as well as security application. As the demand for digital images are increasing, various techniques are established for character recognition. In the character recognition techniques images of documents are taken. Then those images are converted into digital image. From those digital images characters are read and recognized using various techniques.

Optical Character recognition is a promising technology that is used to convert the letters or words written using hand into a digital format. It digitize the printed texts so that they can be electronically edited, searched, stored more compactly, displayed on-line. Optical Character Recognition consists of various stages includes pre-processing, Classification, Post-Acquisition, Pre-Level processing, Segmented Processing, Post-Level processing, Feature Extraction. By making changes in various techniques used in each stages one can get a better result. There were numerous methods were adopted in each stages such as neural networks, fuzzy logic and so on.

This paper gives a literature review on various optical recognition techniques. The next section will give the literature survey of various optical character recognition techniques and final section will give you the conclusion.

II. METHODS FOR CHARACTER RECOGNITION

2.1. Support Vector Machine (SVM)

SVM is usually applied to the classification problem of two classes. For our recognition problem, we need to classify 62 classes plus one class of non-character, which are total 63 classes. Thus, we adopt one-against-one strategy [1] to address the character recognition problem in this paper. Namely, we train one model for any two classes of 63 classes, which produces total 1953 models.

In the testing stage, we apply all trained classifiers to a detection window and then use the voting strategy to estimate the probability of each class which this unknown sub-region. Here, the probability is the normalized voting number.

There are two different features are used to describe the potential characters. Two commonly adopted methods handling this situation are early fusion and late fusion [2]. Early fusion method usually concatenates different features into a longer feature by some fusion strategy. Considering the difference of features adopted in this paper, the late fusion strategy is adopted and the two proposed features are used to train and test dataset independently. For one unknown detection sub-region, a vector formed by 63 probabilities (scores) are estimated for each feature by using above mentioned SVM method. Then we fuse the results of two features, namely add two vectors to get the result vector. Finally, the final label of the unknown region is determined by the component address with the maximum component value of the result vector.

Moftah Elzobi, in his paper, proposed a method for recognizing the arabic characters using Gabor wavelet transform and SVM. It is a segmentation based recognition approach that is used for handwritten arabic letters. This method follows the following steps. Initially, the segmentations of the letters are carried out. Due to the pecularities of arabic characters they have divided the segmentation method into two: i) Resolving the PAW's overlapping and ii) Word image segmentation. To resolve the overlapping connected components and baselines are found and formulated a set of heuristics. In word image segmentation, initially the thinned version is obtained from the overlapping free image. Then a heuristic based election operation is performed. Now feature extraction method is used in the output. In feature extraction Gabor filter based features are extracted. SVM classification is used then to reduce the number of classes. During the training around 1000 word images were used to train the classifier. It is found that recognition rates may vary according to the letters and rate ranges from 43% to 93% was obtained.

2.2. Template Matching

After preprocessing steps, segmented code images are obtained. Template matching can be used for recognizing each code where the template is set to the average of training images. Consider there are C characters to be matched, then the segmented code image is compared with the C templates and it is classified as the code that results in a smallest difference from the corresponding template. As a measure of difference, the sum of squared difference (SSD) can be used which is similar to the variance of pixel values in difference image.

2.3. New Measure

We also have a slightly different variance measure which will be denoted as NM (new measure). Extracting codes from the MRZ and resizing the code image can make small pixel deviation which often occur at the edge of the template image. Thus, the squared difference is divided by 2 if that pixel is on the edge of the template image. This method of New Measure has an effect of reducing high frequency of images.

Let p = {$p1,...,pn$}with $pk(k$ =1...n) be each pixel location and n be the number of pixels. We define D(c)(p)= |I(p) − T(c)(p)|, where I is an input image, and T(c) is a template image of code c. Let E(c) be the Canny edge of each template, then E(c)(pk) is 1 if pk is on the edge, and 0 otherwise.

Finally, the template matching error TM(c) with the template code c is obtained as

$$TM(c) = \sum p \in p\ 1\ E(c)(p)+1[D(c)(p)-D(c)]2, \qquad (1)$$

where D(c) = 1 $n\sum p\in$p D(c)(p). As a result of template matching, the most similar character $c*$ = argmin c TM(c) is found. For E(c) as zero in all pixel, TM(c) is original variance measure. [3]

2.4. Local Binary Pattern

Tasnuva hassan proposed a method to recognize handwritten bangla numerical using local binary pattern. Initial process mainly deals with the noise reduction which improves image quality. This paper uses a method that exploit gausian low pass filter. Next step includes the slant correction. Through this step the slant of writing is corrected. This paper uses KSC algorithm for slant detection and correction. Now all images are normalized to a standard size. For feature extraction of bangla character Local Binary Pattern is used. Three variants of Local Binary Patterns are used. They are basic LBP, uniform LBP, and simplified LBP. After preprocessing, classification of letters is done. This paper uses euclidean nearest neighbor classifier with K=1 is used. The proposed method uses a CMATERdb databases. This data set contains total of 6000 bangla characters. For the test the proposed method uses a 1000 characters that is randomly selected. The performance evaluation is done using different types of LBP features. It is seen that the basic LBP achieves the highest accuracy of 96.7%, uniform LBP sows an accuracy of 96.6%, simplified LBP shows an accuracy of 96.5%.

2.5 Other Techniques

In [4] a work was proposed that will use an open source OCR engine namely Tesseract OCR engine to train a Tamil OCR model. Initial stage in the proposed method is to prepare a training data using computer generated images of tamil characters. The first step in preparing the training data set is to prepare OCR alphabet for the tamil script. As a second step of training, training images are prepared. Next step in character recognition is to segment the character. Character segmentation is performed using the process of creating the bx files. Each character is assigned with a box file, so that the number of boxes is equal to the number of training character. Now training the module is carried out. This proposed method trained the module for individual data set using different images that varies in font size and type. It is found that 3 fonts perform the best. Evaluation process for the proposed OCR system is done using 20 scanned images taken from 20 ancient tamil books and it provides a 14031 characters. It is found that 81% of accuracy is obtained using the proposed system. This is a 12.5% improvement on existing Tesseract tamil module performance [4].

In [5] Ali Farhat proposed a method for automatic recognition of qutari number plates. The input image is converted to binary image. After the initial process of conversion, noise removal plays an important role. Sources of noises are dust particles, scratches in the number plate and so on. After the preprocessing stages, four algorithms have been developed to generate an encoded text from the segmented characters [5]. First algorithm is the vector crossing. It is one of the basic method that is used to perform the OCR. The principle behind the vector crossing is to count the number of vectors passing the character of the image. In order to give a perfect result two vectors one in horizontal and another in vertical are considered. The position of the vectors will be in the center of image. After considering the positions counting starts. Here another method known as zoning comes into picture. It is based on dividing the images into smaller zones and find the density of the character in each zone. Through matching the zones with reference characters to find the best match, the character is identified [4]. The third type of algorithm which uses the advantage of vector crossing and zoning. For example, the density of the image will be calculated and at least one vector crossing will be taken into account. The least vector crossing that is affected by noise, rotation or any other effects was selected to identify the characters. Moreover, sometimes zoning results corresponds to a wrong character due to noise effect; therefore, a specific value was found practically for zoning that a character might deviate from, but in this case, two vector crossings should be satisfied to identify the character [5]. The combination promises the best results. Since the three methods has few disadvantage a fourth method known as Template matching-correlation is proposed. This method finds the correlation between two images. The correlation is evaluated using mathematical equation. The number of reference image is user defined. The four method mentioned above are implemented using MATLAB. As a part of result analysis it is found that that correlation method has the highest recognition rate. The recognition rate is found to be 99.5% and it takes less time to process the image. The lowest recognition rate is for Zoning technique and rate is about 85.3%. The zoning method is affected by the noise. It takes more time to process the image. After making a close study about the recognition rate of crossing and zoning, it is found to be same for both and the rate is about 90.43%. This algorithm combined with the correlation gives a best result. This result is the best result when compared with the accuracy. Zoning recognition rates depends on the character. The recognition rate is best for the number 4 and worst for number 8. However, the noise detection performance of the vector crossing is poor when compared to the existing methods. The rate is found to be 12%. Even though the noise detection performance is poor the recognition rate for the characters is in between 80% to

100%. In order to improve the noise detection performance, it is good to combine both zoning and vector crossing techniques. This leads to an improvement of noise detection rate. The increment in the percent from 12 to 45 is found.

In [6] an OCR system is developed to recognize handwritten kannada letters and numerals. In the preprocessing stage skew detection and correction, binarization, noise removal, normalization and thinning of the input images are carried out. In feature extraction stage both structural features and global transform features are carried out. In order to find the structural features the input image is divided into four quadrant. From each quadrant five features are considered and all of a sum total of 20 features are considered from an image. This 20 features plus one feature namely width feature is considered for an image. The five features that are considered are corner detection, correlation, quadrant density, aspect ratio and width feature. Global features are extracted using Wavelet transform. Wavelets are mathematical functions that cut up data into different frequency components and analyze each component with a resolution matched to its scale [6]. Discrete Wavelet Transform is used in this proposed paper because it is easy to implement and consumes less computational time. In this paper, Symlet wavelet family is used. The symlets are nearly symmetrical wavelets proposed by Daubechies as modifications to the db family and both the wavelet families have same properties [5]. Through this process a set of 128 features are obtained and a set of 21 features are obtained as structural features. A combination of these sets are used in classification phase. Artificial Neural Networks are used in the classification stage. For the training of the system 4350 samples are used and for the testing purpose 1450 testing samples are used. An accuracy of 91% and 97.60% for handwritten kannada characters and numerical are obtained respectively.

In [7] proposed a method which is used to recognize the vehicle number plate using fuzzy logic and scale invariant feature transform. The Scale Invariant Feature Transform or SIFT is a technique which is used to detect, extract and describe local image features. The features extracted are invariant to scaling and rotation, and partially invariant to change in illumination. The obtained features are localized in both spatial and frequency domains. This is done in order to minimize disruption of comparing key features due to noise, occlusion and cluttering [7]. Every images whose features are extracted are subjected to SIFT algorithm. Initial stage of this algorithm is to build a gaussian pyramid. The next stage is to compute the maxima. Last stage of SIFT is to refine the features obtained by calculating several parameters such as the feature's local scale value, the size of feature on original image, flag indicating whether it is an edge or not, the local edge orientation angle and curvature of response through scale space. Mathematically these values are calculated. Features extracted in the SIFT algorithm shall become the values that will appear on the membership function of the fuzzy logic system. Each characters will have its own fuzzy logic algorithm and there will be 7 membership functions in each set of algorithm [6]. The results are analyzed and it is found that 95% of accuracy is obtained on an average.

In [8] a method is proposed to recognize the urdu character using point feature matching. The input image is preprocessed and compare with object templates in the databases. These object stem plate is a combination of single Urdu characters as well as strokes of characters making up a single Urdu word [8]. Once comparison is done the feature points in target images are found. Now the features are extracted from the area of interest and locate the object in the image. It is found that 92% detection is comes up with a positive result.

In [9] developed an OCR system with fast training neural network. In this work each recognition stage is assigned with a training period with short duration. Now the training data are subdivided into several groups. This division is based on the criterion such as symmetry, euler number features and so on. Neural network finds its application during the learning time. This neural network helps the system in learning things at a faster rate. Here in this proposed method no preprocessing is done. Each character is now passed through neural network and selected best one or two results for the next stage. Many OCR techniques make problems in classifying the character in preprocessing stage. This can be overcome by using above mentioned approach. After completing the first stage, the second phase is the similarity check, this will help to omit the similar words. This similarity check is carried out by comparing the pixel of one object with the pixel of another one. Once the recognition results from all neural networks are out, it collects the results, compares it with input character. Now the weighing factor is used to make the final result. The output with largest weighted score is then selected as the final recognition result. In this proposed paper, the recognition rate of this technique is said to have higher precision compared to the conventional neural network approach.

In [10] Optical character recognition is made possible using multilayer perception neural network. As usual the image is aquired initially, then it is preprocessed and segmented. During the segmentation the character lines are seperated. Enumeration of character lines in a character image is essential in delimiting the bounds within which the detection can precede [10]. Next step in segmentation is to separate the characters. Once the characters are seperated the features are extracted. To implement the feature extraction process, Image to matrix mapping process is used. This process is converting the images to a 2D matrix. Next step is to train the system. Training gives the system capability to take the decision to do the task efficiently and it will give a better result in an unpredicted environment. The proposed system used the Multi-Layer Perceptron Learning Algorithm. This method uses pyramid like structure for the learning purposes. This method can be utilized not only for the learning purposes but also for the classification purposes. Appling the learning process algorithm within the multilayer network architecture, the synaptic weights and threshold are update in a way that the classification/recognition task can be performing efficiently [10]. This synaptic weights are important for the iteration purposes. During the iterations the weights are got updated to some integer value. So in order to recognize an object its feature data is feed to the network input layer and produced an output vector. The error is calculated now by the output and by using the target output. By analyzing the output one can determine the character of the recognition rate. The proposed system achieves 91.53% accuracy for the isolated character and 80.65% accuracy for the sentential case character.

In [11] optical character recognition using template matching and back propagation algorithm is implemented. Template matching is one the most common method used in optical character recognition techniques. It is mainly used as a feature extraction technique. Its simplicity for the implementation makes it more popular. Correlation is one other name that holds for the template matching. In this method each individual character pixel matrix is used and they are suitable for the feature extraction. A correlation function R, is used in the test data set and the resultant is the stored in the database. The character with highest correlation value is selected as the best match for that character [10]. Back propogation algorithm that uses reverse mechanism to find the error and it reduces the error by

propagating it backwards. It is based on the error correction. The problem that is found after the grouping. There may be unidentified letters exist after grouping. This unidentified letters will appear as character that leads to an erraneous result. The character recognition using this method gives a highest accuracy rate.

III. METHODS FOR FACE RECOGNITION

3.1. Local Binary Pattern

After the preprocessing step, the Local Binary Pattern (LBP) operator is used for our face description. LBP is one of the best performing texture descriptors and it has been widely used in various applications. It has been proven that highly discriminative and its key advantages, namely, its invariance to monotonic gray-level changes and computational efficiency, make it suitable for demanding image analysis tasks [12]. As originally designed for texture description [13], local binary pattern operator is computed in a pixel level basis using a 3x3 kernel, the surroundings of each pixel can be threshold with the central pixel value, and considering the result as a binary value. The LBP code's decimal form is expressed as:

$$LBP\ (x_c, y_c) = \sum^{N-1}{}_{n=0} f\ (i_n - i_c)2^n \qquad (2)$$

where i_c corresponds to the gray intensity of the center pixel (x_c, y_c), N is the number of sampling points, i_c is the gray intensity of the n-th surrounding pixel, and f(x) is defined as follows:

$$f\ (x) = \{^{0 \quad \text{if } f(x) < 0}_{1 \quad \text{if } f(x) \geq 0} \qquad (3)$$

This basic operator was extended by Ojala et al. [14] to support surrounding points and radius of a pixel neighborhood with different shapes and sizes, enabling handling textures at different scales. Also, the uniform pattern concept was proposed together. A LBP operator is considered uniform if it contains at most two bitwise transitions 0-1 or 1-0 when viewed as a circular bits chain, and Ojala et al. [14] pointed out that nearly 90 percent of LBP operators observed in images are uniform.

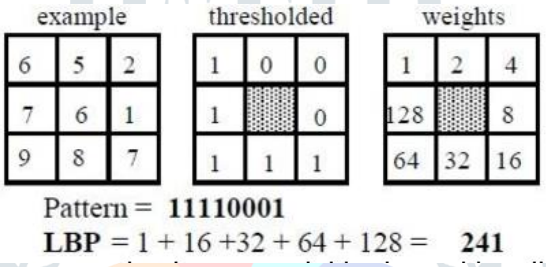

Fig. 1 Basic LBP operation in a 3x3 neighboring, with radius = 1

Fig. 2 Input masked face image (left) and its binary LBP image (right)

It has been shown out that building several local descriptions of the face and combining them into a global descriptor is an effective way of face description [12]. In our work, we used uniform binary based LBP with eight sampling points within two pixel neighborhood. Finally we represent the face region with m sub regions $R_0, R_1, \ldots R_{m-1}$, and compute a histogram independently within each of these m regions. Then we combine the resulting m histograms to yield the spatially enhanced histogram. The spatially enhanced histogram has size (m x n), where n is the length of a single LBP histogram.

3.2. Template Matching

Many studies search for a better way to implement biometric technologies into a system. Matyas and Riha [15] proposed model of biometric verification system. Portal is a step to start authentication process. Input device gets biometric data from a user. Storage is where biometric templates are kept. To get biometric templates from storage, a link between templates and users must exist. Matching Algorithm is an algorithm that verifies biometric data with templates. For accuracy, in this step an appropriate threshold is desired. Result from this unit is either YES or NO. This matching verification process was also supported by Hosseini and Barkhordari [16]. The threshold can define the degree of correlation of each image and is necessary for comparison process. All units work separately while central controlling unit controls all processes and returns verification result [15].

In this research, MPEG7-EHD was chosen to be the descriptor in face matching process. MPEG7-EHD is one of descriptors in MPEG7 visual standard provided by Moving Picture Experts Group (MPEG) used for images and videos. MPEG7-EHD represents the spatial distribution of edges into five categories that are 0°, 45°, 90°, 135°, and non-directional edge. Each image is partitioned into non-overlapping blocks in the same size. Then, five bins corresponding to five edge categories are created to form edge information. This descriptor is invariant to scaling and rotation and also compatible with rotation sensitive matching operations [17].

Verification process requires an appropriate threshold to predict the result of process as mentioned. In this study, the threshold was computed in the same way as [17]. IMG (Rummager) was used as a tool to calculate the distance between two images. Lower distance we get, more similar two images are. Therefore, the low distance between two images can be implied that those images are similar. In the comparison step, there were two parts. The first part was comparison of two images taken from the same person. Such images were a low-resolution frontal face image and a high resolution face image with a selected condition. The average distance d_{sim} corresponding to difference between two images of the same person was calculated. On the contrary, the second part was comparison of low-

resolution frontal face image of a person and high-resolution image with a selected condition of another person. Also, this part computed the average distance d_{dis} according to difference between two face images of different persons. Subsequently, the mid-point d_{mid} of these two distances was found to define as the possible threshold interval. Next, the threshold was increasingly adjusted from d_{mid} - SD to d_{mid} + SD with step size of 0.1SD where SD is the standard deviation of all distances to find out which threshold is most suitable under this situation in terms of false acceptance rate and false reject rate.

IV. CONCLUSION

In this paper we have taken a review of various techniques used for character and face recognition. We have studied about various conditions which an image has to be gone through. Thus by taking an overview of those conditions we can design a technique with high accuracy and high safety. In this paper we have studied that low resolution and low light image may degrade the performance of techniques and must be placed at a proper light place. This paper also concluded that for proper face recognition we need face capture device of proper configuration and light should be proper for high quality image.

REFERENCES

[1] Chang, C.-C. and C.-J. Lin, LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST), 2011. 2(3): p. 27.

[2] Lan, Z.-z., et al., Double Fusion for Multimedia Event Detection, in Advances in Multimedia Modeling, K. Schoeffmann, et al., Editors. 2012, Springer Berlin Heidelberg. p. 173-185.

[3] Hyeogjin Lee, Nojun Kwak, "Character Recognition for the Machine Reader Zone of Electronic Identity Cards," 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, 2015, pp. 387 - 391.

[4] Chamila Liyanage1, Thilini Nadungodage, Ruvan Weerasinghe" Developing a commercial grade Tamil OCR for recognizing font and size independent text" 2015 International Conference on Advances in ICT for Emerging Regions, pp 130-134, IEEE 2015

[5] Ali al-zawqari, abdulhadi al-qahtani, Ali al-zawqari" Ocr based feature extraction and template matching algorithms for qatari number plate" International Conference on Machine Learning and Cybernetics, IEEE 2016

[6] Saleem Pasha, M.C.Padma" Handwritten Kannada Character Recognition using Wavelet Transform and Structural Features" International Conference on Emerging Research in Electronics, Computer Science and Technology, pp 346- 351, IEEE 2015

[7] Rhen Anjerome Bedruz, Edwin Sybingco, Ana Riza Quiros, Aaron Christian Uy, Ryan Rhay Vicerra, Elmer Dadios" Fuzzy Logic Based Vehicular Plate Character Recognition System Using Image Segmentation and ScaleInvariant Feature Transform" 2016 IEEE Region 10 Conference (TENCON) — Proceedings of the International Conference, pp 676-68, IEEE 2016.

[8] Wafa Qaiser Khan, Reema Qaiser Khan" Urdu Optical Character Recognition Technique Using Point Feature Matching; A Generic Approach" Computer Engineering and Technology (ICCET), 2nd International Conference, IEEE 2016

[9] Huei-Yung Lin, Chin-Yu Hsu" Optical Character Recognition with Fast Training Neural Network" pp 1458- 1461, IEEE 2016

[10] Abdullah-al-mamun, Tanjina Alam" An approach to empirical Optical Character recognition paradigm using Multi-Layer Perceptorn Neural Network" 18th International conference on computer and information technology, pp 132- 137, IEEE 2015

[11] Swapnil Desai, Ashima Singh" Optical character recognition using template matching and back propagation algorithm" IEEE 2016

[12] T. Ahonen, B. Hadid and M. Pietikainen, Face Description with Local Binary Patterns: Application to Face Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, pp. 2037-2041, Dec. 2006.

[13] T. Ojala, M. Pietikainen and D. Harwood, A Comparative Study of Texture Measures with Classification Based on Featured Distributions, Pattern Recognition, vol. 29, pp. 51-59, Jan. 1996

[14] T. Ojala, M. Pietikainen and T. Maenpaa, Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, IEEE Trans. Pattern Anal. Mach. Intel., vol. 24, pp.971–987, Aug. 2002

[15] M. Yildiz and M. Gokturk, "Combining Biometric ID Cards and Online Credit Card Transactions," Proceedings of the Fourth International Conference on Digital Society, St. Maarten, pp.20-24, 2010.

[16] Z.Z.Hosseiniand and E.Barkhordari, "Enhancement of security with the help of real time authentication and one time password in ecommerce transactions," in the Proceeding of IEEE 5th Conference on Information and Knowledge Technology., 2013, pp.268-273.

[17] G. Jetsiktat, S. Panthuwadeethorn, and S. Phimolthares, "A Comparison Study of Image Descriptors on Low-Resolution Face Image Verification," In the Proceeding of the 2014 Annual Summit and Conference on Asia-Pacific Signal and Information Processing Association, pp.1-6, Siem Reap, December 2014.