Feature Extraction of Optical Character Recognition Survey from 1980-2010

A. Marial and Jibraiel Jos
Department of Computer Science, Christ University, Bangalore, India

Abstract: Optical Character Recognition is still prevailing even after many decades of implementation. The challenges faced here are increasing day by day so as its applications. From Punched cards to Handwritten Text, from images to video, from uniform font to universal font, from English text to Global language, from researchers to visually handicapped are the transformations obtained from an era of the 1980s to 2010. This study has covered the advancement of acknowledging the characters, how are features are extracted, various methodologies used and more importantly what is the use of OCR.

INTRODUCTION

Acknowledging, Pursuing, Identifying, Interpreting et cetera are the most trivial forms of understandings which vary from person to person. In this process of making to understand it may be easy for Humans at some instances because of their Reasoning skill in natural. But the real Challenges arise when we expect Electrical, Electronic and Digital understandings. This is the reason for the birth of applications for Optical Acknowledgments.

OCR is a field of research in pattern acknowledgment, artificial intelligence and vision. Optical acknowledgment is implemented off-line after the writing either printing has been completed as opposed to on-line acknowledgment where the computer recognizes the characters as they are drawn. Both hands printed and printed characters may be recognized but the performance is directly dependent upon the quality of the input documents. Early optical character acknowledgment may be traced to technology involving telegraphy and creating reading devices for the blind. But its evolution has created the need for identifying text not only from static or still images also including the moving text from videos and day to day applications. Another field of studies has also made it possible to help the visually challenged volunteers to educate themselves independently with no further barriers.

Literature review: It deals with the concept Computer Vision Pattern Acknowledgment (CVPR) compares each algorithm to match with its unique requirements[1]. This paper discussion is based on Text Information Extraction which is most challenging and risk task encountered and the problems which are being mentioned are as follows, firstly Removal of spaces between characters which is direct increases the acknowledgment level. Secondly Comparison between the pre-processed input and existing pattern thus increasing the performance criteria. Thirdly extracting features based on suitable applications. Fourthly performance evaluations, consider any particular algorithm which has its own assumptions and constraints. More importantly, each application uses its own level of complexity, so, it is a tough task of calculating its result. Consists of most of the methodologies used in the field of character acknowledgment in many applications[2]. The study of OCR technology has shown the progress of improvement both for handwritten and printed, so as online or offline characters. This paper takes the steps

involved by OCR technology and has explained all its methods. Instead of taking a black and white form of an input image, this study considers only gray color representation for Initial preprocessing[3]. So as the features for character acknowledgment is extracted from these gray color form of input. It is usually considered that these are more prone to cost. When it is for Handwritten text it is necessary to be in the form of documents and they need to be digitalized but it's, not the same as for Printed text, since, it depends only on printing quality and its paper color.

This study the importance is given to Indian based scripts which are still under research. This includes topics of discussions like Characteristics of each language, suitable methods to use and to find the area of further studies. It also speaks about the latest OCR machines used in the market and its benefits. Features considered in this paper are based on structure and topology of character. The OCR evolution, its necessity, its importance in the field of technology is been discussed. Even though the development of this approach is still growing day by day, its necessity has also increased proportionally. In every field, it has set its footprint. Fukushima and Miyake[4] has the advanced method of identifying patterns with same structure and orientation. It has explained the concepts to replicate the human brain without any external aid. This work is related to neural network considering the eyes with the connection to the nervous system. It is considered to be self-organized system since it is trained but not matched. Identification of text from speech conducted through the phone conversation. Apart from identifying characters, this paper discusses extra features provided like making use of the sound and vocal track to make sure that on the other end it is the authorized person and secondly identifying the person by his voice matching the samples[5].

This study the special task of recognizing words spoken by a human. It is very necessary for the scenarios which need conversations via phone to be traced or tracked. This has many challenges faced since pronunciations may vary from each person to person and the words being used may not always be familiar and understood with the particular context. The most trivial task of work is needed to bifurcate the sound into wanted and unwanted sound. This has a lot of chance of being prone to noise and disturbances. Static image is taken and the characters in it are identified as a concentrated block. Based on matching process, the input block of characters and the existing samples are compared from the repository and a result is favored based on calculations[6]. It involves universal images with different attributes are considered. This study consists of recognizing the text which is been witnessed in the video is very effective[7].

It uses the hybrid wave letter to recognize the text and the sum of squared difference module to track it this process. This takes place frame by frame and it is being matched with the other frames after the identification of the textual and non-textual area. At the end of scanning next the features are being extracted and finally they have the total of 36 features with which the collected data it is been bifurcated into training and testing samples in acknowledgment and the next phase consists of many methods like having multiple resolutions open, obtaining the boundaries for the text and identifying the images with the differences in scale and proportion, the differences in sizes even under Rapid motions.

This study the important concept in OCR technology called as feature extraction is taken and it is made universal for almost all types like printed, handwritten, grayscale, binary and low-resolution character acknowledgment[8]. This is made possible using the concept called as gradient features. The shape of the character in the form of gradient and Direction is based on pixel representation. Thus this paper has shown the maximum acknowledgment rates. The difference between the human acknowledgment of characters and machine acknowledgment is discussed[9]. The advancement in the field of research has emerged many systems with high acknowledgment. This paper has taken the applications of foreign languages and its applications. Mainly the languages like Japanese, Chinese have very high level of complexity because of its structure. Applications using these identified alphabets of expected languages are of great importance in this era of technology. This paper deals with the important challenge of selecting the most appropriate feature extraction methods suitable for any applications[10]. In character recognition based on the extracted features from each character, the performance is evaluated. This paper also deals with various factors which involved in recognition of characters like input character images, properties, disturbances in its different levels, classification types and so on, It explains that character should have following constraints like uniform size and correct orientation, it's handwritten, typed or printed and more importantly it should be known all possible ways in which a character can be written.

The input is in the form of the image from which the standard statistical functions are evaluated and these functions give us the numerical accuracy[11]. Initially, the resultant values are projected in the form of the vector from the output of statistical functions and later based on the differences in value, each character is essentially classified. It explains the complete concepts involved in the path of Optical character acknowledgment. It explains about the difficulties faced in character acknowledgment by taking into considerations of both printed as well as

handwritten text. It even speaks about the importance of OCR with its evolution. It considers the handwritten characters acknowledgment to the network of about >7000 samples[12]. It has the collection of numerals with many handwritings, different size and format. The advantage of this network is that it has highly efficient only if the input sample if with high resolution. The text identification from the image with the different background is the core concept[13]. In this study, the issues related to time and reliability is concentrated. The algorithm follows the importance not only to features extracting but also dividing each letter by letter for further processing is also a major criterion. It takes up the challenge of identifying the character with the acknowledgment rate of 98%[14]. In this paper it takes the input characters in the oldest form of punched cards and based on the concept of Fourier transformations it identifies the parameters for it and it efficiently classified the characters into their respective classes.

Additionally to obtain the text from the image this reference is used[15]. It is explicitly used for the handicapped people with visually challenged. The Training phase contains 117 image samples for blind people and obtained the following observations that Blind people are more prone towards horizontal orientation, clarity of the picture is medium and images are of good quality. The importance is given to classifying the text based on the criteria it satisfies[16]. On the overall view, this concept seems to be simple but this is the major part of character acknowledgment additionally this concept takes a lot of time and energy. The samples are divided based on the number of features it upholds and the present technology uses the machine learning like two classes of samples with a hyperplane, recursive phase of categorizing features, finding the probability among samples and so on. One of the major areas of research which is never ending is the identifying handwritten characters[17]. No two humans have the same type of handwriting if so, then the level of complexity involved in this area is tremendous. When humans find difficult in identifying other handwritings then the real tough task is to make the machines understand. The entire process can be divided into 3 phases of understanding, interpreting and finding it.

OCR design: The brief description about the OCR technology is explained step by step[2]. In the first step (preprocessing) which is the initial step should not have just input alone but based on some criteria, the inputs also should fall under the category of valid and accepted input. The unwanted and inappropriate factors need to be processed in this phase. Characteristics include unfilled segments, nonuniform font, non-defined style, occupying

huge memory and so on. Next follows (segmentation) dividing the entire input into word by word and character by character. This study has given an idea about the combination of two approaches in this step. Image representation being not so important in many of the above papers but here it has been given the good explanation with relevant methods. Identifying change in positions, standard size, considering many resolutions, physical changes, Zonal characters, Measuring distances, features related to ends and edges are highlighted. Once the acknowledgment phase starts so as its usual methods of matching, using methods related to stats, obtaining the traces of structure and mainly concepts related to machine learning are dwelt with. As the last step of combining back the divided characters into words and sentences. The input of characters is taken in the oldest form known as punched cards[14]. When the characters are fed one by one it initially converts the image into matrix form and starts traversing each element of the matrix until it identifies the starting point of the particular character. Character starting point is called to be as Contour, with which the five parameters for the Fourier coefficients are derived.

Initial step the input is taken from zipping format of very large samples thus, the separation of each character from other characters is an initial challenge and then each character is brought under predefined format with specific height and width and preferably not binary image instead gray image is taken[12]. After obtaining the features it is later combined to form High order and then it is fed into the network. Input is taken to be characters of both machine printed as well as handwritten text[3]. So among Initial preprocessing of the input text, this paper says grayscale is the best option than binarization. This is because in the gray scale image the Information loss is prevented but in Binarization since it has a lot of memory required which is subjected to Loss of many information. Accordingly, once gray scale image is taken into account then it consists of Topographical features to be derived from the input. From which an algorithm called as TFG (Topographic Feature Graph) is explained here. This algorithm consists of inputs with the characteristics called as peak, ridge, saddle and flat points along with its other factors like a pit, ravine and hillside. And further, it is converted into GFG (Geometric feature graph) of only points and lines. Thus all these features which are geometric in nature are taken in acknowledgment of characters.

This context consists of 2 kinds of acknowledgment is been explained[9]. In both of these, its own comfort abilities and abilities with some defects can be witnessed. In the first approach, relationships can be clearly defined whereas in second approach the rules defined for that

particular language cannot be saved or upheld. Thus, this study takes the goodness of both approaches and forms a new approach which considers every character uniquely with its own level of characters. This study ahs many concepts discussed which has helped characters at the times of bifurcation of classes. Matching process of obtained input with a predefined pattern, obtaining necessary features from the input, Presenting the input in the form of pixels and obtaining its relationships, the input character is translated into different forms of necessity, Applying the concepts of machine learning, extracting the features using statistical methods. The input is taken as an image, this undergoes initial scanning into the black and white picture and then the features are extracted from the algorithm which extracts the characters from statistical function and further classification with and without error is witnessed[11]. Text extraction from the video is divided into two phases as Scene and Graphic with which this paper concentrates on both these areas[7]. Initially, it uses Hybrid-wavelet to identify the text in the frames of video and further steps of keeping the track of identified text are performed by sum of squared difference based module. In the next step, it uses scale-space feature extractor to remove unwanted gaps witnessed in input image extracted from the text obtained in the video. After that, the authenticated text is being compared with each of the text frames by frame. Finally, the results of the entire result are verified from Contour details obtained from the processed text.

The input taken is a text file with the collection of characters as the input is taken so as the similarity is computed[16]. After which the word counts take place and then follows the extracting necessary features from the machine learning concept and the begins with the classification of different classes. This the initial steps are performed as preprocessing with no big differences. Next follows the extraction of gradient features both in the direction as well as the magnitude and then the feature extraction takes place. In this phase at three levels, the features are extracted mainly binary level, stroke identification and structural level. So, theses values are placed into gradient map form with some degrees and dimensions. Image taken with different backgrounds and different styles are considered and then represented in pixels[13]. From each of these pixels, the pixel which constitutes line both vertical, as well as horizontal line, is taken. Further, the work of making character identified is next step followed by the collection of characters along with the line as a whole the text is rectified and identified. Initially, the input is taken as a collection of frequencies and further is converted into both direction and magnitude removing similar occurrences[18]. Next step follows the traditional method of matching in which the obtained input is matched with a collection of data samples. Based on the outcome the classification phase is conducted.

Input character from the image is taken for pattern determining using the technique of dividing the complete sentence into words, then the words into characters and once the individual characters are obtained, they are taken into gray color representation[6]. Once the defined pattern is appropriate then starts the matching of the obtained pattern with a collection of data samples. The input is from video and image includes both motion and no motion images[1]. It include the variety of letters may be black and white or colored images; it may have movable or nonmovable text in input images. Thus, the problems in identifying text can its own complexity level in different areas of research like Identification of Character, since, it is difficult task to identify required text from the entire image with lot of other elements in it, Determine character exist it includes both based on background changes or various formats of compression, Locating of the character based on resolutions may be both low or high, Extracting text invariant of quality matters and the last one which is only for non-motion images but not for videos which is enhancement of text.

The input is processed to obtain the clear format of necessary character with fewer disturbances and divide the complete characters part by part[17]. Removal of the unwanted stuff involves obtaining the skeleton of characters, applying filtrations, filling the broken edges and so on. The problems faced in all OCR technology is taken into considerations and discussed like, different fonts used, increasing effects of noise, differences in height and width of input. Moreover, if the character is handwritten then it still more tough task with n number of difficulty factors like style, size, orientation and its level of understanding differs from every individual to individual. It has discussed every step right from optical scanning in which it speaks about many scanning devices used, preprocessing which also includes factors like removing of noise, reducing the boundary of text, obtaining skeleton of text and much more. Next follows the feature extraction which involves methods like matching with the pattern, identifying based on pixels representation, classification means all the input has to be divided and categorized into one subclass and last is Post processing.

Feature extraction: This study says that there should be different feature extraction methods and obtains combinations of features from those methods using multiple classifiers and its differences[10]. The best method to find best feature extraction method is using hydrographic map which consists of same orientation,

size and slant. The essential features considered are applying weights to all characters[12]. At any point of time extracting features and relying on features alone constitutes to be similar at the particular point of time. Thus more than the position of the character the more importance is given to the defined weight of each character which may not be the same comparatively to features. The art of extracting necessary components from the words which speak is the really tough task[18]. In this paper, they have discussed the DFT magnitude spectrum which has the highest efficiency in identifying and so the same as Linear prediction which includes many possible samples and at all necessary frequency. When the information is in need of retrieval the available source of obtaining is the voice and along with it words. Emotions expressed along with the pronunciation are another source of information. Since, the data received and information to be produced is very sensitive, this paper has included machine learning concepts and methods.

Many constraints like standard height and width, fixed size and the particular ratio is taken[1]. Pixel values to identify the lines, obtaining the variance of characters linked, refining single component values and finally cluster characters to obtain the entire text. From the basic concepts, this paper gives the prior importance of retaining the shape of any character[17]. Apart from getting many numbers of properties from the undetermined shape, it is advisable to preserve its structure in the process of identification. Additionally, it is also stated that the characters with the single font, uniform shape and definite orientation consist of the very high level of accuracy. The regular structure is advisable in this context. Based on the other end sound and vocal track the words spoken are identified based on some physical features[5]. Every phase consists of weights associated with it so that both aspects of authenticating the other end speaker and along with it identifying the appropriate speaker. Initially, as usual, the input image is prone to scanning and followed by based on Supervised Learning the division of the image into textual and no textual region takes place[7]. As the features are Extracted, It uses Bayes error rate to take decisions of classifying theses features into their respective classes. In this paper, it uses Neural Network Classifier as the features extracted needs to be trained with this and then follows Identifying of the Text from textual and nontextual regions. Next follows the Tracking of the text. In this paper, the tracking follows three categories of text included like static, simple linear motion and complex nonlinear motion which efficiently tracks all these three classes. Methods used for recognizing are as follows Matching with many Resolutions of images, minimum SSD, Text Contour and finally Analyzer of Motion.

It has briefly examined methods used for feature extraction[2]. Taken from identifying the blocks of characters to bringing back to identified words, it involves n number of features like Size, Location, Orientation, layouts, change in position, change in angle, Curves and

points, Extreme points, Maxima and Minima, points ,cross, ballpoints, junction, lines and many more. When the image is said to be with the presence of some characters then it has to be proved with some conditions[6]. This study marks out some of the conditions satisfied like Color differences which can be witnessed based on background and the particular text content, to make sure that the characters are not left unidentified the next condition it requires if equal size for all characters and along with the size it also gives importance to uniform light and reflection. With the strict guidelines of all these above conditions, the present misinterpretations and misclassifications will be easily suppressed. Features of input character are either matched with samples taken as standard or input characters are taken as pixel pieces of information[9]. Once these features are extracted then they are taken into row and column consideration of representation. Additionally, the input characters are subjected to change of place and change of orientation under some specifications. Physical characters line lines, closed sections, the number of ends and still more features are added in this technique. Fourier method of transformation being used to obtain essential characteristics from each character by character and later compared with patterns[14]. Then follows the comparison between the sample and the pattern to categorize into their respective classes. The characters with different characteristics and transformations like differences in scales, differences in orientation and differences in shape and size are taken care. In fourier representation, it considers the degrees of rotational constants up to three. The images with good quality, good clarity and of horizontal orientation are taken[15]. AdaBoost is the algorithm followed in this paper which helps in obtaining the classifiers to be effective but one factor in this is that the images should be labeled as either text or nontext. Thus, it suggests obtaining only essential and useful features instead of obtaining all available features.

It takes into considerations of text identified from the video and images are taken for acknowledgment[1]. Again here the challenges for every input in character acknowledgment are doubled. Since, characters are not of the expected format and especially in expected proportion. So as in this paper, it includes the entire cross verifications performed for such characters with the challenging task of Identification of characters in motion. Both in the video as well as in Images the challenges it faces includes variations of light may be contrast or dark colors, Motion captured both in moving position or in constant still position, Texture of letters may differ in style, shape and fonts and most importantly disturbances produced by characters from other factors. It uses the geometric features for extracting features, firstly Identifying the Outline, here it obtains the two levels of derivatives then follows the calculations of Eigen values and Eigen vectors and the finishes with labels, secondly Obtaining initial information here it obtains the values of

four points like peak, ridge, saddle and flat with which TFG is evaluated[3]. Thirdly Extracting elements, in this, the obtained TFG is lead to the construction of GFG and lastly Identification under study. This paper, it describes all methods involved in the process of feature extraction. Among all the other phases of OCR, this is the most challenging task. In this paper it explains about Pattern being matched with all respective characters, representation of the character into pixel representation, Vectors based on character shape, Density based acknowledgment and still many features extracted when it has differences in shape, size, orientation and the font is considered.

As the system, if initially trained with few samples and it addresses few features and the next time when it encounters similar features those pattern acknowledgment can be effortlessly witnessed[4]. It is not just one or two rather numerous patterns need to be used to train the system. The works of so far studied OCR has obtained the level of accuracy and acknowledgment rate. On the side of improvements, few more points included are breaking the barrier of the quality of material like paper, documents and scanner. Development of the inputs to OCR from selected font styles to all possible fonts. Retaining complete information of the character rather than limited features, Most necessary field of Handwritten needs to be enhanced and the most important among all is that OCR being available and helpful to physical handicapped blind people. It considers the updated version of OCR technology which not only extracts the essential characteristics from an alphabet but more than that it also reduces the burden of error during the classification. Features related to size, shape and internal structure additionally features related to the topology of characters. For all the languages there is unique way of identification propose like in Devnagari, the characters are firstly divided into a number of parts and again it is further divided until precise measure and performs further work. In Bangla, the entire technique is based on the relationship between the structure of characters. In Tamil, the features taken into considerations are based on a number of rounds, directions, angles and wavelet characters. In Telugu it uses the typical pattern matching technique, in Oriya the features can be extracted directly without the need of preprocessing, in Gurumukhi the features like endpoints, junctions, loops and positions are taken care, In Gujrathi all concepts related to techniques of Machine learning are taken, normalin Knormal OCR steps like scanning, identifying features and further processing with machine learning concepts.

Gives the algorithm taken the example as human face detection which extracts the defined characters like, input should be of the single person, shots with different angles, different expressions, involving all possible motion and more importantly having different views[11]. But this also

has on drawback as the background should be same for all shots taken. Input consists of 60images out of which the rate of classification without error, for the algorithm defined is about 87.5% to 100%. It takes the second sample of some machinery parts taken are 220 inputs and its rate of classification has given 100%. The use of gradient representation has made the processing as universal and its rate is high comparatively[8]. Here, it keeps the height and width of the input image to be at the constant value but not for handwritten. It also solves one of the most challenging tasks of identifying similar characters of different classes is solved based on contours. For every image, Gradient is calculated and a margin is set to remove the unwanted features. Here gradient values obtained are put into gradient map. The presence of pixels, stroke, variants features are obtained from this gradient map. To make it simple initially features based on selected criteria is not considered and only with limited characters are taken for further steps[16]. Finding the similarity, considering only 2 classes at a time classifying features recursively until it cannot be further divided, obtaining the probability and lastly defining the margin and decision based on the margin are the methods used to obtain characters and to classify (Table 1 and 2).

Use of OCR: Various methods of obtaining the acknowledgment of characters which come out of the human's mouth[18]. Since, this study has covered almost all the possible and most importantly all the efficient methods with the highest acknowledgment rate. It has taken into-considerations of possible error at times like noisy background, non-matching words with the database, non-sink between speaker and listener, noise, interruptions and still many. Many of the concepts explained this paper presents Support Vector Machine, the machine learning concept to be the most efficient one[16]. It consists of border line drawn between the two classes considered as two groups of data. This is more suitable and this step avoids the confusion faced in classification. Additionally, this evolves to be faster than the other concepts of categorizing (Fig. 1).

This study it speaks about the advancement of technology in OCR[19]. Earlier OCR was based only on matching the input samples with constrained samples. But nowadays OCR not only matches but also performs the evaluation by selecting the necessary features. Thus, it is clear that the present OCR has become universal with all types of inputs. The development trend which the character acknowledgment has set is very neatly explained. In many fields of area OCR has emerged to be more efficient but in this paper, its challenges and methods to overcome the challenges are briefly discussed. Fixed fonts, Fixed size, Isolated characters, noiseless background, printed characters and still much more are the criteria required for the acknowledgment rate to be

Table 1: Comparison of methodologies used

Nature of paper	Recognition rate of success	Methods used
[10] Survey (Isolated characters)		1. Template matching 2. Deformable templates 3. Unitary image transforms 4. Graph descriptions 5. Projection histograms 6. Contour profiles 7. Zoning 8. Geometric moment invariants 9. Zernike moments 10. Spline curve approximation 11. Fourier descriptors.
[14] Implementation (handwritten)	98%	1. No optimized decision method (Fourier transactions for Feature Extraction)
[1] Survey (Text in video and image)		Text information Extraction(TIE) (TIE includes a. Text detection b. Text Localization c. Tracking d. Extraction and e. Enhancements)
[7] Implementation (Text in video)	Text = 88% Non text = 77%	1.HYBRID WAVELET (Neural network based method) 2. Sum of Squared Differences based (Tracking module uses a. SSD based Image Matching b. CONTOUR based stabilization)
[3] Implementation (machine printed and handwritten)		1. Matched filters (grayscale) 2. TFG topographic feature graph 3. GFG geometric feature graph (Topographic features like peak points, ridge points, saddle points and flat points, pit, ravine and hillside)
Survey (machine printed and handwritten)		All process involved in OCR technology. (1. Optical scanning 2. Preprocessing 3. Feature extraction 4 classification 5 post processing)
[8] Implementation (Isolated characters) for all like printed, handwritten, grayscale, binary and low resolution	Handwritten and machine printed is 98 and 99.4% Machine printed 97.5% Low resolution 88.2%	Gradient-based contour 1. (Gradient computation 2. Feature computation 3. Stroke detection)
[18] Implementation (Speech recognition)		1. Vector Quantization 2. Gaussian mixture model 3. Support vector machine 4. Fusion (a. Selection of features (b. Speaker modeling (c. Feature extraction (d. Speaker modeling (e. Robust speaker recognition (f. Super vector methods
[19] Survey (Indian language scripts)	Devnagari-96.5 and 97.5% Bangla- 96.8% Gurumukhi-97.3% Oriya- 96.3%	Studies on Devnagari, Bangla, Tamil, Oriya, Gurmukhi, Gujarati, And Kannada character recognition
[9] Review paper		1. Template matching and correlation techniques 2. Feature analysis and matching
[11] Implementation (Image recognition)	100% (220 samples)	Optimal discriminant criterion (Minimum distance classifiers)
[16] Implementation (Natural language text)	92%	Inductive learning methods 1.(Find similar 2. Decision tree 3. Naive Bayes 4. Bayes nets 5. Support Vector Machine)

Table 1: Continue

Nature of paper	Recognition rate of success	Methods used
[1] Implementation (natural scene)	90.39%	Stroke width Transform
[5] Implementation (speech identification)	96.8%	Gaussian Mixture Speaker
[2] Review (handwritten)	95%	All process involved in OCR technology. (1. Optical scanning 2. Preprocessing 3. Feature extraction 4. Classification 5. Post processing)
[6] Implementation (Image)	93.4%	Character pattern candidate
[17] Survey (handwritten)		a. Online handwriting recognition b. Offline handwriting recognition 1. Preprocessing 2. Character recognition 3. Word recognition
[15] Implementation (Image)	90%	AdaBoost

Table 2: Comparison of factors necessary for OCR

Paper index	Factors considered	Measurements for character classification scheme
[10]	Translated, stretched, scaled, rotated, skewed and mirrored character original shape	Character original shape
[14]	Size, rotation and arbitrary phase angle	Statistical mean, Probability and distance measurements
[1]	Size, alignment, intercharacter distance, color, motion and edge-compression	Word by word identification, character by character identification and pixel by pixel classification
[7]	Font size, font style, poses, shapes, colors and text under complex motion, reserve text and normal text	Frame by frame, static, scrolling and Complex motion
[3]	Features (for grayscale images) like Peak, Pit, Saddle, Ridge, Ravine, Flat, Hillside, Ridgeline, Ravine line, flat line and hillside line	1. TFG topographic feature graph 2. GFG geometric feature graph
Impedovo, etc	Fonts, noise, translation, rotation, scaling, NLP-derived differences in size and pitch	Based on result from Feature extraction phase, the classification phase determines classification result of text
[8]	Gradient features, binary features, stroke features, gradient map	1 low level 2 high level 3 structural level
[18]	1. Physical differences (vocal tract shape, larynx sizes, parts of voice production) 2. Manner of speaking (accent, rhythm, intonation style, pronunciation pattern, choice of vocabulary)	Matching Input signal with database
[9]	Location, background, Multifont	1. Global transformation and series expansion 2. Features from statistical distribution of points 3. Geometric and topological features
[16]	Size, NLP-derived phrases, binary features	Factoids, multi-word dictionary entries, noun phrases
[1]	Stroke width transform	Each pixel having stroke
[5]	Frequency warping, spectral shape compensation, Weights, mean, variance	
[2]	Slant angle, Contour direction, zoning, Topological features, Maxima and minima, Reference lines, ascenders and descenders	1. Whole word/ sub-word recognition 2. Limited / unlimited vocabulary 3. Long words 4. Segmentation errors 5. Explicit/implicit segmentation
[6]	Size, font and gray level	1. Character candidate region 2. Character pattern candidate 3. Relaxation operation
[17]	Continuous line, dotted line, and curvilinear velocity	
[15]	Mean, entropy, scale, orientation and Font	

high but when these conditions are not met then it gives direct impact on efficiency. In [19] it follows the new approach of continuous training leads the system to take up the individual decisions. So as the system is trained with many samples, it will gain the ability to identify the patterns automatically. As in [2], It is useful to identify letters which are not easily identified or classified. Additionally based on obtaining Fourier parameters it is very efficient to divide the characters into categories of classes with efficient preprocessing of input characters. As the use of grayscale is enhanced here and so as the scope of improvement is also doubled up when the input images with the lower resolution are also taken into considerations[3]. More precisely the memory and processing time has drastically reduced. This technique has increased the acknowledgment rate. The concept of weight has played the vital role[12]. As per neural network algorithm, here with thousands of data samples is been related to output with the minimum effort put for initial processing of data. The network used here is
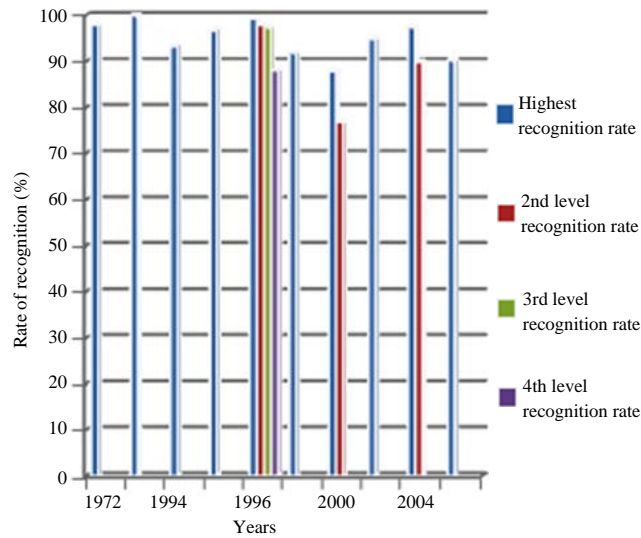
Fig. 1: Evolution of OCR with the Recognition rate of characters across the years from 1980s to 2010

consisting of some predefined constraints like redundancy and time. The output of this network is further applied and implemented. Images with the characters are not easy to be identified by systems, it faces the major challenge in the clarity and the interpretations of the machine are highly challenging[6]. The initial differences between the background of the image and the area which has the characters are found out. For the universal characters with no standard size, font and shape this paper has given the acknowledgment rate to be 85.7%. The similarity in structure and shape is important factor described here. OCR technology started from Isolated text from punched cards and now it has evolved in identifying handwritten characters of different fonts and style. From all these developments and the studies many factors like acknowledgment rate, accuracy has increased. From printed text to handwritten text methods and approaches used may differ but the ultimate goal of character acknowledgment has been the standard and unique goal.

Present methodologies of recognizing characters are discussed[9]. Nowaday's the characters identifying devices are available in small handheld devices taking care of factor of comfortableness. As the requirement increases so as the efficiency also needs to be considered. To increase the efficiency rate it also considers some factors like Input characters should be isolated, clear background, position and font types. The acknowledgment rate of text from images is about 99% up to 0.73 precision values with 0.94 seconds[1]. This paper uses both pixel calculations as well as local values of character to obtain the result 15 minutes faster than the present system. When the characters even with heavy background are identified this is the way towards perfection. Consists of not only to reduces the processing time each frame by frame but more

importantly[7]. It upholds the uniqueness of this application is to maintain the Acknowledgment even when additionally data are included constantly and continuously. With the use of obtaining starting point of the character in each frame and with additional it reduces Error in the textual region from Multiple frames.

Attains the acknowledgment rate of 96.5% in 5 sec of time duration for speech[5]. Over telephone, it's 80.8% of acknowledgment rate in 15 seconds with 49 samples. The Gaussian method used based on sound and vocal track work well even with heavy background and less time. Comparatively the performance, has increased drastically and moreover it is in related to all levels of the input sound is taken. Along with all the other features it is not just selected features but instead, the combination of features including structure and semantic features provides the essential features[2]. In the matching process, it seems to be pretty well and easy obtainable but it is heavily disturbed with the other criteria's. This paper also suggests that following all the steps involved by OCR is not enough to obtain the high performance instead effective methods need to be adopted. Including concepts from Machine learning and Artificial Intelligence so that with the information from different sources will be an aid to improve the system. This it has universal solutions for all the problems being faced in OCR[8]. All the necessary information is obtained from the one important factor called gradient map which uses both pixel representations as well as outline representation also. With which not only the acknowledgment rates have been increased but including results for handwritten has also increased proportionally. Hence, this method is most flexible and it can be made still more efficient as research goes on in this field.

It includes an algorithm which uses the statistical functions and comparatively with the present system, this algorithm obtains more benefits like high accuracy rate, improved acknowledgment, any type of input with no barrier of clarity and dimension[11]. It also suggests the further improvements in this field as reducing the classes of classification and different criteria are given to input. Address interpretation of the handwritten text, many bank documents verifications, signature and writer identifier and much more[17]. The urge of digitalization has opened the facilities of handwritten characters acknowledgment. Apart from this concept, it can also be used as the method of improvement when the error in handwriting is identified and rectified. This paper[1] consists of extracting the characters both from the video that is the movable and images that are no movable inputs. So, it is necessary for each of the text that is been extracted from this inputs has to be put into the OCR technology. By which the necessary features can be obtained from the input into the requires text format. When initially the input image is fed, output. OCR performs the initial preprocessing for the input image in terms of conversion of the input image into binary text format and further preprocessing steps is being performed in order to get the output. It not only includes the extraction of images from the video as well as the image it also includes the process of identifying its characters, locating its positions apart from the noise encountered. Additionally, this paper drastically decreases the overall processing time and removes the maximum challenges faced in obtaining those characters is also taken into account. The key role played is to enable to identify only essential features[15]. This paper has obtained the accuracy of 90% which is comparatively good. Moreover, this has also made way for the classifiers to be strong from weak classifiers.

CONCLUSION

The current situation with the workmanship in OCR has moved from primitive plans for constrained character sets to the use of more refined procedures for Omni-text style and imprint acknowledgment. The primary issues in OCR more often than not lie in the division of debased images which are joined or divided. By and large, the precision of an OCR framework is specifically reliant upon the nature of the input. Three figures are utilized as a part of appraisals of OCR frameworks are right grouping rate, dismissal rate and mistake rate. The execution ought to be evaluated from the frameworks blunder rate as these mistakes pass by undetected by the framework and must be physically situated for correction.

Despite the greater number of calculations that have been created for character acknowledgment, the issue is not yet understood satisfactorily, particularly not in the situations when there are no strict restrictions on the handwritten or nature of print. Up to now, no acknowledgment calculation may the rival man in quality. In any case, as the OCR machine can read much speedier, it is still advancing.

REFERENCES

01. Epshtein, B., E. Ofek and Y. Wexler, 2010. Detecting text in natural scenes with stroke width transform. Proceeding of the 2010 IEEE Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, IEEE, Redmond, Washington, ISBN: 978-1-4244-6984-0, pp: 2963-2970.

02. Arica, N. and F.T. Yarman-Vural, 2001. An overview of character recognition focused on off-line handwriting. IEEE Trans. Syst. Man Cybernet. C Appl. Rev., 31: 216-233.

03. Wang, L. and T. Pavlidis, 1993. Direct gray-scale extraction of features for character recognition. IEEE. Trans. Pattern Anal. Mach. Intell., 15: 1053-1067.

04. Fukushima, K. and S. Miyake, 1982. Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition. In: Competition and Cooperation in Neural Nets, Amari, S. and M.A. Arbib (Eds.)., Springer, Berlin, Germany, pp: 267-285.

05. Reynolds, D.A. and R.?. Rose, 1995. Robust text-independent speaker identification using gaussian mixture speaker models. IEEE Trans. Speech Audio Process., 3: 72-83.

06. Ohya, J., A. Shio and S. Akamatsu, 1994. Recognizing characters in scene images. IEEE. Trans. Pattern Anal. Mach. Intell., 16: 214-220.

07. Li, H., D. Doermann and O. Kia, 2000. Automatic text detection and tracking in digital video. IEEE Trans. Image Process., 9: 147-156.

08. Srikantan, G., S.W. Lam and S.N. Srihari, 1996. Gradient-based contour encoding for character recognition. Pattern Recognit., 29: 1147-1160.

09. Govindan, V.K. and A.P. Shivaprasad, 1990. Character recognition-a review. Pattern Recognit., 23: 671-683.

10. Cilibrasi, R. and P.M. Vitanyi, 2005. Clustering by compression. IEEE. Trans. Inf. Theory, 51: 1523-1545.

11. Liu, K., Y.Q. Cheng and J.Y. Yang, 1993. Algebraic feature extraction for image recognition based on an optimal discriminant criterion. Pattern Recognit., 26: 903-911.

12. LeCun, Y., B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard and L.D. Jackel, 1989. Backpropagation applied to handwritten zip code recognition. Neural Comp., 1: 541-551.

13. Epshtein, B., E. Ofek and Y. Wexler, 2010. Detecting text in natural scenes with stroke width transform. Proceeding of the 2010 IEEE Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, IEEE, Redmond, Washington, ISBN: 978-1-4244-6984-0, pp: 2963-2970.

14. Granlund, G.H., 1972. Fourier preprocessing for hand print character recognition. IEEE. Trans. Comput., 100: 195-201.

15. Chen, X. and A.L. Yuille, 2004. Detecting and reading text in natural scenes. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2004 Vol. 2, June 27-July 2, 2004, IEEE, Washington, USA., pp: II-366-II-373.

16. Dumais, S.T., J. Platt, D. Heckerman and M. Sahami, 1998. Inductive learning algorithms and representations for text categorization. Proceedings of the 7th ACM International Conference on Information and Knowledge Management, November 2-7, 1998, Bethesda, MD., pp: 148-155.

17. Plamondon, R. and S.N. Srihari, 2000. On-line and off-line handwriting recognition: A comprehensive survey. 1EEE Trans. Pattern Anal. Mach., 22: 63-84.

18. Kinnunen, T. and H. Li, 2010. An overview of text-independent speaker recognition: From features to supervectors. Speech Commun., 52: 12-40.

19. Pal, U. and B.B. Chaudhuri, 2004. Indian script character recognition: A survey. Pattern Recognit., 37: 1887-1899.