ISSN: 2277-9655
Impact Factor: 4.116

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

PUNJABI CHARACTER RECOGNITION USING WAVELET AND CURVLET TECHNIQUES WITH SIFT ALGORITHM

Er. Balwinder Singh*, Er. Surekha Rani
* Student of M-Tech in Ludhaina College of Engineering and Technology Ludhaina
Assistant Professor Ludhaina College of Engineering and Technology Ludhaina

DOI: 10.5281/zenodo.56025

ABSTRACT

There are many techniques used to recognise the characters and out of these one is named as curve let technique and other is wavelet technique. The Punjabi characters has different type of patterns as compared to English characters. Curve let is a non- adaptive technique for multi scale object representation and curved singularities can be well approximated with very few coefficients and remain coherent waveforms under the action of the wave equation in a smooth medium. Standard wave let transforms for two dimensional functions have only very crude capabilities to resolve directional features. The usual orthogonal wave let transforms have wave lets with primarily vertical, primarily horizontal and primarily diagonal orientation. The handwritten characters have different orientations. Therefore the handwritten recognition of characters is challenging research area in the field of pattern recognition. Wavelets generalize the Fourier transform by using both location and spatial frequency. For 2D or 3D signals, directional wavelet transforms go further, by using basis functions that are also localized in orientation. A curve let transform differs from other directional wavelet transforms in that the degree of localization in orientation varies with scale. In this paper we are using both curve let and wavelet techniques to recognise the Punjabi characters and compare the results.

KEYWORDS: Wave let, Curve let, Recognition,Pattern.

INTRODUCTION

Optical character recognition (OCR) is process of automatic computer recognition of characters in optically scanned and digitized pages of text. OCR is one of the most fascinating and challenging areas of pattern recognition with various practical application potentials. It can contribute immensely to the advancement of an automation process and can improve the interface between man and machine in many applications[1].The renewed interest in on-line handwriting recognition stems from a number of factors. Compared to the 1960's,now have more accurate electronic tablets, more compact and powerful computers, and better recognition algorithms. However , there are additional and perhaps more important reasons. First, the recent hardware advance of combining tablets and flat displays bring input and output into the same surface. This combination permits the use of electronic ink, providing immediate feedback to the writer of the digitized writing[2]. The off-line recognition is dedicated to bank check processing, mail sorting, reading of commercial forms, etc, while the on-line recognition is mainly dedicated to pen computing industry and security domains such as signature verification and author authentication[3]. Curve lets provide a new multi resolution representation with several features that set them apart from existing representations such as wavelets, multi wavelets, steerable pyramids, and so on. Handwriting recognition allows the creativity of handwriting to be combined with the advantages of word processors[4]. The more relevant patterns at your disposal, the better your decisions will be. This is hopeful news to proponents of artificial intelligence, since computers can surely be taught to recognize patterns[5]. Devanagari script is a two-dimensional composition of symbols attached to the top, bottom, left or right of a main character. There are approximately 93 symbols in Devanagari of which there are 13 vowels, 36 consonants, 28 pure consonants (half-forms) and 16 modifier symbols besides a number of conjunct forms. Many other symbols are derivable from these. Since the script composition in Devanagari is two dimensional and the number of symbols is large, its input using a keyboard is cumbersome. This makes the online pen computing environment very attractive

ISSN: 2277-9655
Impact Factor: 4.116

for complex scripts like Devanagari[6]. In image analysis applications, the situation is different. When working with real rather than synthetic data, one of course doesn't `know' where these edges are; one only has a digitized pixel array, with potential imperfections caused by noise, by blurring, and of course by the unnatural pixelization of the underlying continuous scene. Hence the typical image analyst only has recourse to representations which don't `know' about the existence and geometry of the discontinuities in the image[7]. Standard wavelet transforms for two-dimensional functions have only very crude capabilities to resolve directional features. The usual orthogonal wavelet transforms have wavelets with primarily vertical, primarily horizontal and primarily diagonal orientations[8]. Mainly, character recognition machine takes the raw data that further implements the process of pre-processing of any recognition system[9]. The internal factors are distortion in the character images during scanning of images addition of noise during image acquisition degraded and broken characters images. These factors are responsible for the reduction in the accuracy of the handwritten character recognition[10].

CHARACTER PROCESSING

There are basically two techniques are used in Character Processing i.e. Wavelet Technique and Curve let Technique

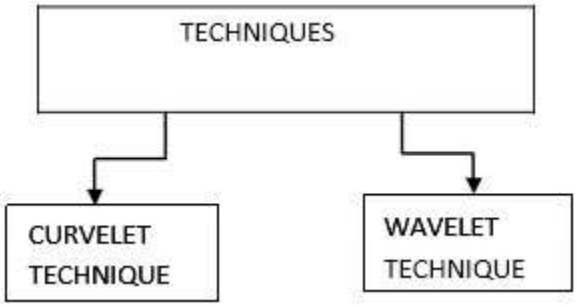

Fig 1:- Character Processing Techniques

Curve let technique:- Curve lets provide a new multi resolution representation with several features that set them apart from existing representations such as wavelets, multi wavelets, steerable pyramids, and so on. They are based on an anisotropic notation of scaling. The frame elements exhibit very high direction sensitivity and are highly anisotropic. The curve let construction was originally developed for providing efficient representations of smooth objects with discontinuities along curves; the underlying motivation being to apply this construction to classical image processing problems such as Data compression, compression of digitally acquired image ,Image restoration, image reconstruction or edge-preserving regularization.

Wavelets technique:- Wavelets generalize the Fourier transform by using both location and spatial frequency. For 2D or 3D signals, directional wavelet transforms go further, by using basis functions that are also localized in orientation. A curve let transform differs from other directional wavelet transforms in that the degree of localization in orientation varies with scale. Wavelets are Related classical multi resolution methods use a Limited dictionary made up of roughly isotropic elements occurring at all scales and locations.

Image processing: In this process, we apply different algorithms to get smooth image. The different operations i.e. binarization, noise reduction, skew detection and skeltonization of character image will use for pre-processed image. In binarization process, we convert image in to bit pattern. Noise reduction algorithm will apply on binarized image to remove unwanted bit-pattern, which do not have any significance in the output.

Enhancement: Image enhancement is basically improving the interpretability or perception of information in images for human viewers and providing `better' input for other automated image processing techniques. The principal objective of image enhancement is to modify attributes of an image to make it more suitable for a given task and a specific observer.

ISSN: 2277-9655
Impact Factor: 4.116

Normalization: Text in handwritten images typically shows strong variability in appearance due to different writing styles. Appearance differs in the size of the words, slant, skew and stroke thickness. Such variability calls for the development of normalization and pre-processing techniques suitable for recognition of handwritten text.

Curve let Transformation and Wave let Transformation: Curve let is a non- adaptive technique for multi scale object representation. Being an extension of the wavelet concept, they are becoming popular in similar fields, namely in image processing and scientific computing.Wavelets generalize the Fourier transform by using both location and spatial frequency. For 2D or 3D signals, directional

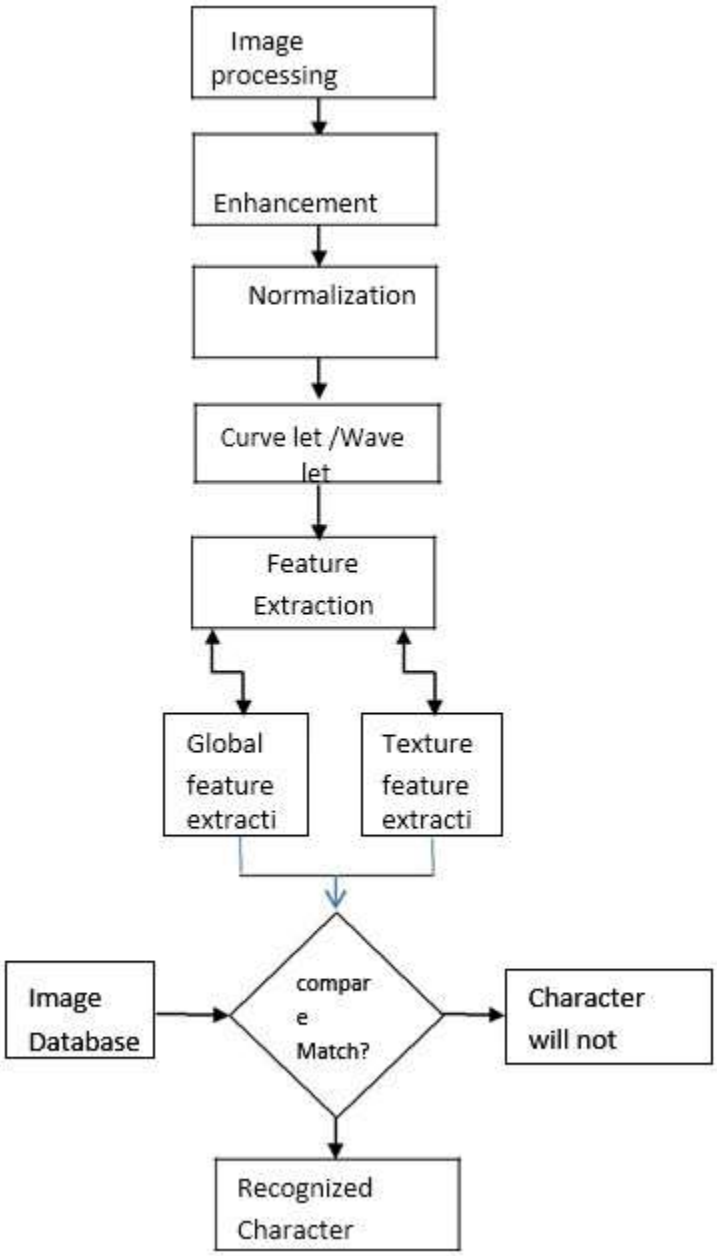

Fig 2:- Block diagram of character recognition using curve let and wave let

wavelet transforms go further, by using basis functions that are also localized in orientation. A curve let transform

differs from other directional wavelet transforms in that the degree of localization in orientation varies with scale. Wavelets and related classical multi resolution methods use a limited dictionary made up of roughly isotropic elements occurring at all scales and locations.

Feature extraction: Each character has some features, which play an important role in pattern recognition. Punjabi characters have many particular features. Feature extraction describes the relevant shape information contained in a pattern so that the task of classifying the pattern is made easy by a formal procedure.

Image database: The starting point of the project was the creation of a database with all the character images that would be used for training and testing. The image database can have different formats. Character images are handwritten digitized images. This means that characters on paper have different sizes, different resolutions and sometimes almost completely different angles. Images belonging to the last case was very few but they were discarded, as there was no chance of classifying them correctly.

Matching: After extraction features from image, then we apply matching algorithm to image. For this process, referred image will match with template image. In matching algorithm, match the features (that were extracted in step 2) of template image to referred image.

RESULTS

The results of employing the Punjabi character matching method are very promising. The visual results of applying the algorithm to an image obtained by a common optical scanner. To verify the results, we create a database by using optical scanner. We acquire all Punjabi characters of different lipi at different environment conditions. The method also performs surprisingly well on different Punjabi Character images acquired at low resolution images. With low resolution images, there is more noise and as a result, more unique locations. Table 1 demonstrates the Genuine Acceptance Rate of applying the algorithm to an image taken from the database.

Aside from the basic visual results, the algorithm output was tested in two ways, first for accuracy, and secondly for reproducibility.In this test, the method achieved better than 91% accuracy, our experiments show that features carry significant discriminatory information. There is a relative decrease of 20 percent in the equal error rate (EER) of the matching system. This significant performance gain is consistently experimental observed across various quality images.

Genuine Acceptance Rate		
Threshold value	High Resolution image	Low Resolution images
45	91	75

CONCLUSION

This Paper has proposed a method for matching a Punjabi character features. We have initially found that, based on anatomical evidence, but further analysis based upon an analysis of different kinds of characters of different kinds of lipis. Along with that, it has through discussion of cross correlation algorithm.

This approach will give the technology new amplitude in order to provide a secure way of recognising Punjabi Characters, in which the unique features are logically extracted. This proposed algorithm performs better than existing recognition algorithms.

ISSN: 2277-9655
Impact Factor: 4.116

REFERENCES

[1] U. Pal and B. B. Chaudhuri, "Indian script character recognition", Computer Vision And Pattern Recognition Unit, Indian Satastical Institute Colkata Vol. 37(9), pp. 1887-1899 (2004). C. Y. Suen, C.C. Tappert and T. Wakahara,
[2] "The State of the Art in On-Line Handwriting Recognition", IEEE Trancisions on Pattern Analysis and Machine Intelligence Vol. 12, No. 8(1990).
[3] F. Bortolozzi, A. Britto Jr., L. S. Oliveria and M. Morita, "Recent advances in handwriting recognition", in the Proceedings of International Workshop on Document Analysis (IWDA), India, pp. 1-30 (2005).
[4] T. Wakahara, H. Murase and K. Odaka, "On-line handwriting recognition", Proceedings of the IEEE, Vol. 80(7), pp. 1181-1194 (1992).
[5] A. K. Jain, R. P. W. Duin and J. Mao, "Statistical pattern recognition: a review", IEEE Transactions on PAMI, Vol. 22(1), pp. 4-37 (2000).
[6] S. D. Connell, R. M. K. Sinha and A. K. Jain, "Recognition of unconstrained on-line Devanagari characters", in the Proceedings of 15th International Conference on Pattern Recognition Vol. 2, pp. 368-371 (2000).
[7] E. J. Candes and D. L. Donoho. " Curvelets: A surprisingly effective non-adaptive representation for objects with edges", Vanderbilt University Press, Nashville, TN pp. 1-10 (2006).
[8] E. J. Candes and D. L. Donoho "Continuous curvelet transform: I. Resolution of the wave front set" vol. 19, p.162-197 (2006).
[9] G. Singla and Dr. P.Kumar"Extract the Punjabi Word from Machine Printed Document Images", Int. Journal of Engineering Research and Application Vol. 3, Issue 5, pp.343-348(2013).
[10] P. Singh & S. Budhiraja "Handwritten Gurmukhi Character Recognition Using Wavelet Transforms", IJECIERD ,Vol.2, Issue 3 pp.27-37 (2012).