ISSN: 0974-2115
www.jchps.com Journal of Chemical and Pharmaceutical Sciences

An approach to multi secure to historical malformed documents using integer ripple transfiguration

K.P. Kaliyamurthie*, P.C. Balasubramanian
Dept. of CSE, Bharath Uni versity, Chennai, Tamil Nadu, India.
***Corresponding author: E-Mail: kpkaliyamurthie@gmail.com**

ABSTRACT

Document Analysis and Recognition (DAR). The general objective of DAR research is to fully automate the process of character recognition, which can be performed on handwritten or printed text of good quality or distorted documents. Distortion finds a common occurrence in any place of written or printed document. After careful analysis of many samples, it is identified that there are several kinds of distortions seen in almost every scripts used in any part of the world. Hence, there arises the need to cope with these problems to make a secure document process for the recovery of distorted text.

Currently, there is no software available for securing distorted scripts in general and distorted Tamil scripts in particular. The entire process consists of three main stages. The first stage is the segmentation stage which takes the input as an image of a document and separates the different logical parts The second stage is feature extraction stage, which analyses the set of distorted document based on set of features that can be used to uniquely identify the document. The final stage is the classification stage which is considered as the major decision making stage of the entire system. Recognition accuracy is achieved by increasing the PSNR (Peak Signal Noise Ratio) value. In addition document is going to be secure in order to avoid further distortion. Entire process is going to give value addition in the field of DAR especially for ancient Tamil Script

KEY WORDS: DAR, PSNR, Pattern Recognition, Character Recognition.

1. INTRODUCTION

India is famous for its rich cultural heritage even before prehistoric period. Life style of Indians, their behavior, attitude, faith in poetic justice, hospitality, religious faith, their literature were known from various sources, such as inscriptions on temple walls, rocks, pillars, engravings in caves, paintings, copper plates, literature in palm leaves and reports of foreign travelers etc. Among them, perishable materials like palm leaves become the prey for insects, poor maintenance, handling, ageing, fire, etc. Few materials, created thousands of years before, though available in good form, are written in old scripts and unable to recognize easily.

Many a great men dedicated their life with devotion and dedication in collecting the rare scripts all over the country and published them with great difficulty. Scholars with good experience, inspiration and well versed in this field are capable of studying the scripts and compare them with similar materials. They have the ability to extract the right sense from the scripts. Resolving the ambiguities is another problem. Optical Character Recognition (OCR) is one of the most fascinating and challenging areas of Pattern Recognition (PR) with various practical applications. Pattern Recognition is the ability to extract the structure of an object of study with the help of applied inputs and prepared specimens put away as references. It can contribute tremendously to the headway of a computerization prepare and can enhance the interface in the middle of man and machine in numerous applications. One such functional application in PR is the field of Document Analysis and Recogn(DAR).

DAR is limitless and contains numerous applications. Character Recognition (CR) frameworks are a subset of Pattern Recognition. Upon this, numerous examination works are in advancement comprehensively. Design acknowledgment puts a crucial part in determining the aforementioned issues. Design acknowledgment examination is finished just on the off chance that it perceives misshaped report, which is inadequate in the present exploration works. Edge identification is one of the basic strides in picture preparing, picture investigation, picture design acknowledgment, and PC vision procedures. In later past example acknowledgment specifically bargains PC vision frameworks, introduction and force data about edges as essential info for further handling to report recognizable proof. The edge can have diverse importance in different connections. As needs be, diverse edge identification calculations can perceive edges in distinctive types of representation and each of them can be considered as an authentic edge discovery calculation. Edge diminishing calculation alongside Sobel is taken care of to learn the examination work. Edge identification calculation helps us to find and distinguish sharp discontinuities from a picture that finds the edges in an info picture by approximating the inclination extent of the picture.

In the present challenge the archives identified necessities mystery and protection. Steganography allows a client to conceal a lot of data inside of picture and sound documents. These types of Steganography frequently are utilized as a part of conjunction with cryptography so that the data is much ensured; first it is scrambled and afterward concealed so that an enemy needs to first discover the data and after that unscramble it. Steganography is utilized to conceal the data with the goal that mystery can be set up. Cryptography is utilized to scramble the data sent and make it as a mixed content which is not justifiable. In the event that Steganography and Cryptography are utilized together it has more point of preference for giving security to the dat

In the present challenge the archives recognized necessities mystery and security. Steganography allows a client to shroud a lot of data inside of picture and sound records. These types of Steganography regularly are utilized as a part of conjunction with cryptography so that the data is profoundly secured; first it is scrambled and after that concealed so that an enemy needs to first discover the data and afterward decode it. Steganography is utilized to conceal the data with the goal that mystery can be built up. Cryptography is utilized to scramble the data sent and make it as a mixed content which is not reasonable. On the off chance that Steganography and Cryptography are utilized together it has more favorable position for giving security to the data The framework of the project is to retrieve the text from a distorted document which distorted during Electronically Shared Documents like photocopier, scanning, faxing, digital image capturing using etc., Proposed system focuses to rectify the electronically distortions caused by page curl, perspective, border noise due to bounded volume books and their overlapping of text. Any form of distorted input undergoes adaptive thresholding to binarize the input image and retrieve a set of connected component regions (CCs).

Ordinary archive acknowledgment in caught pictures utilizing computerized camera. Viewpoint bending is one of the primary difficulties for acknowledgment in caught pictures subsequent to the camera might regularly not have a front-parallel perspective of the content. This system for point of view recuperation of content in characteristic scenes, where content can show up as separated words, short sentences or little sections (as found on publications, announcements, shop and road signs and so on.). It depends on the geometry of the characters themselves to evaluate a correcting homograph for each line of content, independent of the perspective of the content over a vast scope of introductions. The exactness of today's report acknowledgment calculations come up short suddenly when archive picture quality contorted somewhat

The level viewpoint fitting so as to foreshorten is redressed two lines to the top and base of the content, while the vertical point of view performing so as to foreshorten and shearing are assessed a straight relapse on the shear variety of the individual characters inside of the content line.

Digitization is a technique by which the gray scale pictures are changed over to paired pictures. In any picture investigation or improvement issue, it is exceptionally crucial to recognize the objects of enthusiasm from the rest. Digitization isolates the closer view (message) and foundation data. This is exceptionally valuable for the acknowledgment of twofold sided mutilation. Digitization is normally answered to be performed either all around or locally. Worldwide strategies apply one force quality to the whole picture. Neighborhood or versatile edge systems apply diverse power qualities to distinctive areas of the picture. These edge qualities are controlled by the area of the pixel to which the limit is being connected. A few digitization strategies are examined. In our exploration range we are going to take after both Global and Local digitization technique utilizing Image securing.

Pre-Processing: Main objective of this pre-process is to remove all variations that may affect the final identification results. This involves operations on the digitized image intended to reduce noise and increase the ease of extracting process include noise removal, de-skewing (Skew detection and correction), line removal, symbol removal, skeletonization, etc. Except skeletonization, we have selected some of the existing operations to pre-process the script. Salt and pepper noise have been removed using standard algorithm given by Iliescu et al.. It has been noted that there is always some extent of skewness present in the document image that may lead to the poor performance of an OCR system.

Segmentation: It is one of the essential steps in any OCR. Segmentation is the process of segmenting the whole document image into separates text from graphics, tables, pictures etc. These approaches perform segmentation based on general features such as white space and pitch. Content part from the file, It may contains many lines, is extracted and the segmentation step is followed. Further, each line is then segmented into individual words and finally each word has to be segmented into its constituent alphabets. It most commonly employed to extract the lines from the document. Similarly a vertical projection of a line image can be used in the extraction from the word. The research community generally ignores character segmentation, even though broken alphabets can possible of majority of errors in automatic reading of both machine-printed and handwritten texts.

The feature extractor generates a set of features that enables the classification to distinguish between patterns of various characters. The feature extraction selects a set of feature which will uniquely identify the individual character. This selection set of features is the heart of pattern recognition system design. Various types of features that have been used are edges, closed loops, strokes etc. of diminished result. This procedure can uproot anomalies Sobel method proves to perform well regardless of the intensity differences between foreground and background they just need to work on one pixel wide. It additionally decreases the information records and undoubtedly.

2. CONCLUSION

This paper aimed to develop a secure distortion less document image. The focus was on using the Sobel method to detect edges and to secure data with IWT. The thesis demonstrated a set of novel ideas and approaches which use Sobel method with the global structure to detect the distorted exploring a large area in order to overcome

distorted. To handle the data in a secure manner, Steganography are carried out using reconciling thrashing algorithm. Resultant document evidences remarkable hiding rates with Steganography systems.

REFERENCES

Achudhan M, Prem Jayakumar M, Mathematical modeling and control of an electrically-heated catalyst, International Journal of Applied Engineering Research, 9 (23), 2014, 23013.

Anderson RJ, FAP Petitcolas, On the limits of steganography, IEEE Journal on Selected Areas in Communications 16 (4), 474–481.

Gopalakrishnan K, Sundeep Aanand J, Udayakumar R, Electrical properties of doped azopolyester, Middle - East Journal of Scientific Research, 20 (11), 2014, 1402-1412.

Gopinath S, Sundararaj M, Elangovan S, Rathakrishnan E, Mixing characteristics of elliptical and rectangular subsonic jets with swirling co-flow, International Journal of Turbo and Jet Engines, 32 (1), 2015, 73-83.

Ilayaraja K, Ambica A, Spatial distribution of groundwater quality between injambakkam-thiruvanmyiur areas, south east coast of India, Nature Environment and Pollution Technology, 14 (4), 2015, 771-776

Kerana Hanirex D, Kaliyamurthie KP, Kumaravel A, Analysis of improved tdtr algorithm for mining frequent itemsets using dengue virus type 1 dataset: A combined approach, International Journal of Pharma and Bio Sciences, 6 (2), 2015, 288-295.

Lingeswaran, Prasad Karamchet SS, Gopikrishnan M, Ramu G, Preparation and characterization of chemical bath deposited cds thin film for solar cell, Middle - East Journal of Scientific Research, 20 (7), 2014, 812-814.

Premkumar S, Ramu G, Gunasekaran S, Baskar D, Solar industrial process heating associated with thermal energy storage for feed water heating, Middle - East Journal of Scientific Research, 20 (11), 2014, 1686-1688.

Sundar Raj M, Saravanan T, Srinivasan V, Design of silicon-carbide based cascaded multilevel inverter, Middle - East Journal of Scientific Research, 20 (12), 2014, 1785-1791.

Thooyamani KP, Khanaa V, Udayakumar R, Application of pattern recognition for farsi license plate recognition, Middle - East Journal of Scientific Research, 18 (12), 2013, 1768-1774.

Thooyamani KP, Khanaa V, Udayakumar R, Efficiently measuring denial of service attacks using appropriate metrics, Middle - East Journal of Scientific Research, 20 (12), 2014, 2464-2470.

Thooyamani KP, Khanaa V, Udayakumar R, Partial encryption and partial inference control based disclosure in effective cost cloud, Middle - East Journal of Scientific Research, 20 (12), 2014, 2456-2459.

Thooyamani KP, Khanaa V, Udayakumar R, Using integrated circuits with low power multi bit flip-flops in different approch, Middle - East Journal of Scientific Research, 20 (12), 2014, 2586-2593.

Thooyamani KP, Khanaa V, Udayakumar R, Virtual instrumentation based process of agriculture by automation, Middle - East Journal of Scientific Research, 20 (12) 2014, 2604-2612.

Thooyamani KP, Khanaa V, Udayakumar R, Wide area wireless networks-IETF, Middle - East Journal of Scientific Research, 20 (12), 2014, 2042-2046.

Udayakumar R, Kaliyamurthie KP, Khanaa, Thooyamani KP, Data mining a boon: Predictive system for university topper women in academia, World Applied Sciences Journal, 29 (14), 2014, 86-90.