

Development of Braille and Latex Markup Code for Mathematical Equations using OCR

Kavyashree M K
Department of ECE
SJCE, Mysuru, Karnataka, India

Shraddha S K
Department of ECE
SJCE, Mysuru, Karnataka, India

Ranganath K N
Department of ECE, SJCE, Mysuru,
Karnataka, India

Nitish B M
Department of ECE
SJCE, Mysuru, Karnataka, India

Sirisha S
Department of ECE
SJCE, Mysuru, Karnataka, India

Abstract— This project demonstrates a system that takes a photograph of a printed equation and produces a Braille and Latex code representation. The process uses adaptive thresholding morphological edge smoothing, and the Hough transform for image binarization and skew correction. The project uses Hue invariant moments and circular topology to match characters against a stored database of characters. The algorithm then assembles the appropriate Latex code from the detected characters and then it is converted into Braille. First, the image is captured. The target user does this through a Smartphone camera. Then, the image is binarized to obtain clean image. Then, skew is corrected so the equation in the image is horizontal. Afterwards, segmentation algorithm using the bounding box, edge detection and convex hull find each character in the equation, extract a feature vector for each, and identify the characters using nearest neighbor classification. Finally, our developed algorithm assembles the recognized characters into LATEX code and then it is converted into Braille. This system is able to detect 90% of characters in ideal images and 75%-85% of characters in real world photographs.

Keywords—Binarization; Skew Correction; Hough Transform; Bounding Box.

I. INTRODUCTION

Detection of text in equations and identification of characters in scene images containing mathematical equations is a challenging process. And also LaTeX is a powerful typesetting system that is extremely useful mathematical equations. However, once rendered, the output cannot be modified with-out access to the code. Re-coding lengthy equations is time consuming. Taking a photograph of an existing equation printed in a textbook and to produce editable LaTeX code solves this problem. And also Braille output of the corresponding mathematical equation helps the blind to take up research activities in the field of Mathematics and in other fields where mathematics is involved. The steps to achieve this involve converting the photograph to a binary image, correcting skew, segmenting characters, matching the characters to a previously stored database of characters,

generating the correct LaTeX representation of the equation and producing braille output of the same

II. LITERATURE SURVEY

An approach for reducing morphological operator dataset and recognize optical character based on significant features [2]. Pattern Matching is useful for recognizing character in a digital image. Optical Character Recognition (OCR) is one such technique which reads character from a digital image and recognizes them. Line segmentation is initially used for identifying character in an image and later refined by morphological operations like binarization, erosion, thinning etc. We considered this paper for details on binarization. An integrated skew detection and correction using fast Fourier transform and dct [8]. Skew detection and correction is very important If it not detected correctly it will lead wrong result in future during image analysis. To measure the processed time and speed taken by skew detection algorithm. This technique is used by firstly applying DCT compression and thresholding on image to reduce timing computation and after that Fourier spectrum is obtained. Further this spectrum is divided into four quadrants and detected skewed angle of each quadrant is measured. And finally Input image is rotated by using bilinear interpolation method. We have used Hough transform instead of Fast fourier transform. Translation, Rotation, and Scale-Invariant object recognition [5] Invariant-object recognition (IOR), whose aim is to identify an object independently of its position (translated or rotated) and size (larger or smaller), has been the object of an intense and thorough study. In the last several years, an increasing number of research groups have proposed a great variety of IOR methods. Among them, we can find a number of optical techniques, boundary-based analysis via Fourier descriptors, neural-network models invariant moments, and genetic algorithms. Visual pattern recognition by moment invariants [9] In this paper a theory of two-dimensional moment invariants for planar geometric figures is presented. A fundamental theorem is established to relate such moment invariants to the well known algebraic invariants. Complete systems of moment invariants under translation, similitude and orthogonal transformations are

derived. Some moment invariants under general two-dimensional linear transformations are also included.

III. PROPOSED SYSTEM

A. Binarization

The input RGB image is first converted into a binary image for processing. Treating scanned Images differently from Smartphone photographs gives the best results. In order to differentiate between the two image types, an image is first converted to grayscale.

B. Skew Correction

We then correct the binaries image for rotations. To compute the dominant orientation, we take Horizontal Profiling..

C. Segmentation

Because characters are matched individually, the characters are first extracted using Segmentation.

D. Matching

Each character found by the segmentation algorithm through a nearest neighbor classifier using Manhattan distance, which produces the best results compared to other types of distances tested and best match is the detected character.

E. Equation Assembly

We assembly each equation sequentially from left to right using each recognized character's bonding box. And then convert into Latex and Braille

IV. IMPLEMENTATION

A. Binarization

The input RGB image is first converted into a binary image. In order to differentiate between the two images types i.e scanned and photographed image, image is first converted to grayscale. Then, we find the proportion of pixels that are mid-gray range, defined as having an 8-bit intensity value between 15 and 240. If this proportion is less than 0.1, the image is classified as a scanned image otherwise, as a photographed image. Photographs taken by a Smartphone contains uneven lighting conditions. So we use adaptive thresholding with noise removal to compensate.

B. Skew Correction

Then the binarized is corrected for rotations. To calculate the dominant orientation, we take the Hough transform. Fortunately, most equations have multiple horizontal lines, such as fraction bars, equal signs. This means that the dominant orientation is usually given by the horizontal orientation.

C. Segmentation

The characters are first extracted from the skew correction algorithm output. Centroids and bounding boxes of edge maps used as a segmentation algorithm because of the ability to extract characters surrounded by others. The edge map is obtained by eroding the inverted image and then XORing that with the original inverted image. Then, for each edge, we extract its centroid, bounding box and convex hull.

D. Matching

Each character found by the segmentation algorithm is matched through a nearest neighbor classifier using Manhattan distance.

E. Equation Assembly

Equation assembly is done from left to right using each recognized character's bounding box. As we process each character, we keep track of a previous centroid to detect the presence of superscripts and subscripts. If the current character's bounding box does not overlap with next bounding boxes, we add the character to the equation directly. For a limit-enabled control sequence like summation, integration, and product we recursively assemble the equations of the upper and bottom limits as appropriate.

V. RESULT ANALYSIS

We tested our system for various inputs, both clean and smart-phone photographed images, observed the process flow and the results obtained and have documented the same. The original image selected is as show in fig 1. The fig 1 is then binarized, deskewed, segmented, matched against a previously stored database and then finally converted to Latex and Braille Format as shown in fig 2 and fig 3.

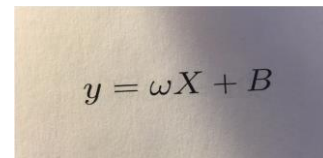


Fig. 1. Original Image

$y = \omega X + B$

Fig. 2. Latex output



Fig. 3. Braille Output

VI. APPLICATIONS

Converting mathematical equations in textbooks form image form to braille helps the blind to take up research activities in the field of Mathematics and in other fields where mathematics is involved.

LaTeX is a powerful system that is extremely useful for technical documents. If a graduate student or professor wants to reproduce existing equations from textbooks, technical papers etc, in latex form, Recoding a lengthy equation is time consuming and prone to error. This project makes this process easy by converting photograph of a printed equation to the LaTeX code to reproduce the equation.

VII. CONCLUSION AND FUTURE WORK

The primary goal of this project is to develop a system that can detect a mathematical equation from a photograph and generate LaTeX markup code and producing braille output for the same. We have implemented this project in six modules namely, Binarization, Skew Correction, Segmentation, Character Identification, Character Matching

and Equation Assembly. Each of these modules is implemented and tested accordingly. Our system is distinguishable from other systems in a way that it is cost effective and directly focuses on the detection and extraction of characters which is more accurate. Relevant future additions to our "Equation Assembly" algorithm include the ability to detect and extract an equation from a page with background text and clutter. The current algorithm only supports equations with completely empty backgrounds. This project has also been architected to be readily extended to recognize handwritten equations. We chose the features for character matching to be invariant to scale and rotation to support the huge variances in handwriting. Further work can also be done to increase the accuracy of the current matching algorithm. The matching algorithm exhibits susceptibility to binarized and smoothed characters that were not perfect PDF images.

REFERENCES

- [1] A. Pradhan, M. Pradhan, and A. Prasad, "An approach for reducing morphological operator dataset and recognize optical character based on significant features," in Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on, vol. 2, Aug. 2015, pp. 1631-1638..
- [2] C. Singh, N. hatia, A. Kaur, "Hough transform based fast skew detection and accurate skew correction methods", International Journal of Scientific and Technology, vol. 41, no. 112, 2008, pp. 3528-3546.
- [3] M. K. Hu, "Visual pattern recognition by moment invariants", IRE Trans. Inform. Theory, vol. IT-8, 1962, pp. 179-187.
- [4] M. Kaur and S. Jindal, "An integrated skew detection and correction using fastfourier transform and dct", International Journal of Scientific and Technology Res, vol. 2, Dec. 2013.
- [5] M. K. Hu, "Visual pattern recognition by moment invariants," IRE Trans. Info. Theory, vol. IT-8, 1962 pp. 179-187..
- [6] J. Ashley, R. Barber, M. Flickner, J. Hafner, D. Lee, W. Niblack, D. Petkovic, "Automatic and Semiautomatic Methods for Image Annotation and Retrieval in QBIC", SPIE Proc. Storage and Retrieval for Image and Video Databases, 2015, pp.24-35.
- [7] R. Sarkar, S. Malakar, N. Das, S. Basu, M. Kundu and M. Nasipuri "Word Extraction and Character Segmentation from Text Lines of Unconstrained Handwritten Bangla Document Images", IET Image Processing, vol 1, 2011, pp. 227-260.
- [8] O. Arandjelovic, R. Cipolla, "Automatic cast listing in feature-length films with anisotropic manifold space", Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, vol 2, 2006, pp. 1513-1520.
- [9] C.C. Tappert, C.Y. Suen, T. Wakahara, "The State of the Art in Online Handwriting Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 8, August 1990.
- [10] L.Chandrasekar,G.Durga"Implementation of Hough Transform for Image Processing applications Communications and Signal Processing (ICCSP),2014InternationalConference"Vol.56,No.3,2014,pp.962-967.
- [11] Bruguera J.D., Guil N., Lang T., Villalba J. and Zapata E.L., "Cordic based parallel/pipelined architecture for the Hough transform," VLSI Signal Process., Vol. 12, No. 3, 2007, pp. 207-221.
- [12] R. E. Schapire, "The boosting approach to machine learning: An overview" in Nonlinear estimation and classification, New York:Springer, 2003, pp.149-171.