Automatic Logo Extraction and Detection for Document Verification using SIFT and SURF

Vaijinath V. Bhosle
Computer Science,

College of Computer Science & Information Technology
(COCSIT), Latur, Maharashtra, India,

Dr. Vrushsen P. Pawar
Science Faculty
Professor and head in Water and Land Management Institute
(WALMI), Aurangabad, Maharashtra, India,

Abstract— **Algorithms and methods used for recognize of objects from images in imperative to implement numerous operations such as document verification and object identification. To extract and segmenting logo from transfer certificates or any other document become the difficult task. The present investigation is effort to design the necessary programme module to extract logo from different document.**

This well-defined module could be an efficient tool to recognize the valued document like TC or any other document that enhances accuracy to saves time , this Automated logo detection module and its application could be extended in access of other documentation. In this paper SIFT and SURF algorithm is used to extract logo from image. In the next step work may be used in noise removal technology. This module and its proper execution help to extract and matches the objects and will give promising results for document identification. We got 96% accuracy.

Keywords— *TC-Transfer Certificate, Logo Detection, Thresholding, feature extraction, SIFT, SURF, Bounding Box, match feature.*

I. INTRODUCTION

Preset investigation is effort to recognize and verify the valued document by using automated logo detection module that basically involves segmentation of objects, extract its features, match the particular features with the unique documents. Document like transfer certificate or any other document is an unique Logo of particular institution that is not supposed to be copied or made duplicated that has huge consequences with academics, present study assures not only correct identification of transfer certificate by its logo but also will enhances the purity of valued documents.

Next step up of present study is to extend this tool to implement its application in access confidential document like stamp papers, farmer valued document, government certificates etc.

Logo is unique for different institute. It could be find the logo is belong to which institute. Thus, each origination has his one unique logo, but this unique logo change in size, color, texture and pattern.

The SIFT or SURF is used to extract the feature from image and match the feature for verification, it gives beater result compression with other algorithms. Many researches are used SIFT and SURF function to extract or match the feature of different images. Therefore, there is a need to figure out a simple but effective process module to be used. [3]

Carneiro and Jepson is work on logo detection as per his opinion to perform the operation by using local image features extraction.[1]

This research is concerned with study and analysis of focusing document image retrieval and logo matching in journal database. [2]

The main goal of the proposed module is to extract the logo from different document of different institute. Match the logo with document is it valid or not. To find out it is original or duplicate. It will automatic verify by the module.

II. LITERATURE SURVEY

In 2013 Fanman Meng, Hongliang Li, [8] they presented work on web logo detection by using MRF method in used to segment the logo from images and then SIFT method is used to match the feature with logo as per his opinion it gives good result.

In 2014, Remya Ramachandran, Andrews Jose [3] presented logo matching and recognition system using SURF algorithm. Using SURF descriptors extract the interest points from reference logo and test image. Then implement a model free context similarity function to find the similarity measure. Using the similarity measure find the matching between reference logo and test image. The performance is evaluated and compare the results with SIFT descriptors. SURF algorithm is better than the SIFT algorithm and will give better matching.

In 2014, Prof. Mrunalinee Patole, Meera Sambhaji Sawalkar,[4] they presented novel solution for logo matching and recognition based on Context-Dependent Similarity (CDS) kernel is proposed and it's able to match and recognize multiple instances of multiple reference logos in image archives. Query logo and target logo images consist of spatial context of local features like interest points, regions.

In 2015, B.Sai Prasanthi, D. Rajesh Setty[5] they extracted six feature from images by using sum technique. This mechanism involves a amount of components with image processing, edge detection, image segmentation, characteristic extraction, comparing images. The characteristics extraction is performed on the image of the

currency and it is compared with the characteristics of the genuine currency. The Sobel operator is used to perform this operation. It is batter other than other operator. Paper currency recognition with good accuracy and high processing speed has great importance for banking system as well as in other sector. MATLAB is batter software to extract the feature from different images. In this paper MATLAB is used to detection of paper currency. It gives batter result for extraction.

In 2015, Mustafa S. Kadhm, Lim Shu Yun [6] in this paper they develop module framework is used to logo detection of Vehicle Brand by using OpenCV function. This is complete tool for detecting Vehicle Brand of logo or as part for image noise removal technology it is an very important in computer vision applications. By developing this framework it is batter to test and match the logo it gives very nice result for detecting logo Vehicle Brand.

2015, Mrs.Prajkata Musale, Prof.Mr. Deepak Gupta,[7] they presented the operation on high motion videos for logo matching and recognition. The central issues of this technology are fast localization and accurate matching and unveil the malicious use of logos that have small variation with respect to the original. This paper introduce, matches and multiple instances of multiple archives quotes with equality logo to identify a variation framework to conceive of innovative mastering assistance through quotes logo and test images Planetarium localized symptoms like interest points are reflected at least one agreed by mixing the power function[8].

III. METHODS

A. SIFT Algorithm

SIFT algorithm is used for to match the image on the interest point detection. This is used for extract interest point feature from the image. That is used to match logo with different document. This algorithm is work on the bases of four steps 1)Scale- space construction, keypoint localization, orientation assignment and keypoint descriptor.

In the first steps find out location and scales of key points in the Difference- of Gaussian function.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma) \times I(x, y) \ldots (1)$$

In the second steps rejected the low contrast feature from image.

In the third steps key point of local image obtained on the bases of gradient.

In the fourth stapes compare each key point on the bases of centered key point. with 4×4 array location grid and 8 orientation bins in each sample. That is 128-element dimension of key point descriptor[9].

Fig. 1 Images After SIFT Algorithm

B. SURF Algorithm

SIFT is used for extract the feature from image. This algorithm is very slow to overcome the limitations of SIFT, other algorithm SURF is used to extract the features very fast. Speeded up robust features technique (SURF) used integral images for image difficulties and Fast-Hessian detector. This is very fine as compare to SIFT. This algorithm creates a "stack" with the same resolution. The box filtering is used for to filter the stack SURF[10].
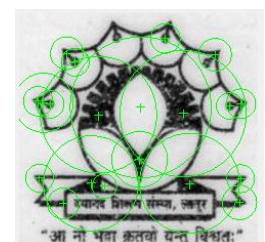

Fi.g 2 Original image   Detected SURF features on image
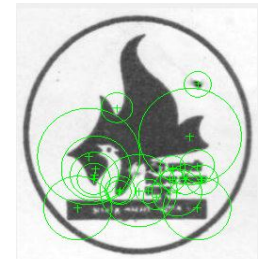

Fig. 3 Original image   Detected SURF features on image

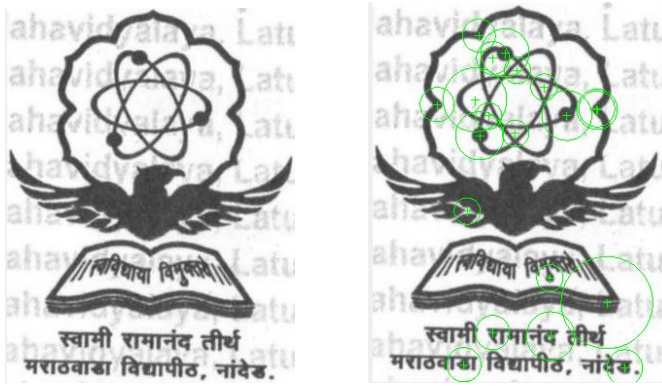Fig. 4 Original image   Detected SURF features on image

Fig. 5 Original image   Detected SURF features on image

TABLE I

Features Display on Each Image using SURF (in this table first five features are displayed but for the experiment 20 features are took place)

Img/Feat	Feature1	Feature2	Feature3	Feature4	Feature5
Image 1	343.9	343.7	176.8	182.4	76.7
Image 2	119.2	47.1	160.5	118.2	217.1
Image 3	192.5	192.0	197.1	84.3	76.7
Image 4	110.5	259.9	136.7	66.9	200.3

.

IV.   SEGMENTATION

In computer system, image is divided different segment this process is known as segmentation. The area of segmentation is to simplify the problem for document verification. This paper can use different method for segmentation Threshold based segmentation, Edge based segmentation, and Region based segmentation.

Regions consistent to separate surfaces of images depend on classification of different objects present in images. Segmentation is used for object extraction from document images for document analysis.

V.   PROPOSED SYSTEM

A.   Algorithm

Step1 : Read Image from database x=image.

Step2 : Store it into variable.

Step3 : if size(I,3)==3 then.

Step4 : Convert RGB into grayscale.

Step5 : Convert Grayscale image to binary image by

Thresholding

Step6 : Extent regions properties of image.

Step7 : Apply Bounding Box function on Document

Images

Step8 : Extract Objects from Transfer Certificate

B.   Flow-chart of the proposed Document Logo Detection
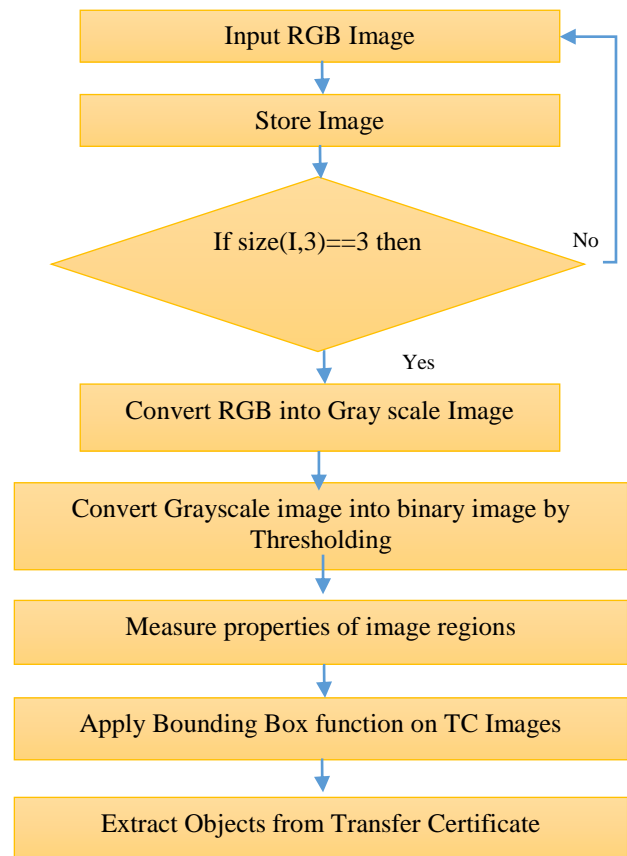
C.   Motivation for the Logo Segmentation

There are two main areas in this work. The first is to generate the database for logo of different organization and second is to extract the logo from transfer certificate develop segmentation algorithms for images of Transfer Certificate or any other document, in which color and pattern typically do not exhibit uniform statistical characteristics. The second is to incorporate knowledge of human perception in the design of underlying feature extraction algorithms.

Segmentation of images of Transfer Certificate is particularly difficult because, unlike artificial images that are composed of more or less pure textures, the texture properties are not well defined. The texture property of perceptually different regions are not even due to effects of lighting, perspective, scale changes, etc.

Number equations consecutively. Equation numbers, within

D. Image Database

This paper generates its own data base that is related to logo of different institution for document verification. This database is consisting of 100 logo images from different institutions.

E. Extracting Logo from TC

This method uses document image for the process of verification. In this process document image is provided to the system for different operation. In this module RGB image

is given to the system. Convert the image RGB to gray. Extract the logo by using min max method. Apply threshold method to normalize the image. Bounding box function is used to select the region form image. Sobel edge detector is used to detect the edges from logo for matching operation.

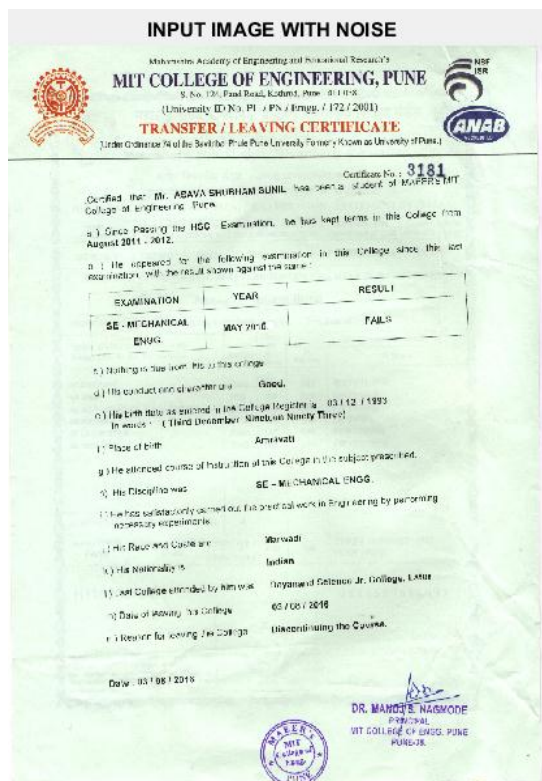Step 1 Input RGB Image

Fig. 6 Original RGB image

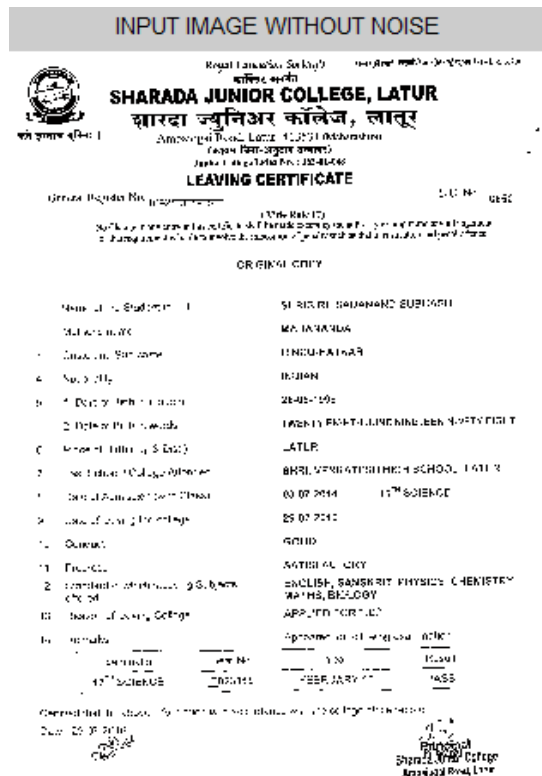Fig. 7 Original RGB image

Step 2 Convert RGB to Gray Scal Image.

Fig .8 RGB to gray

Step 3 Extent region of image properties.

Step 4 Select particular logo from image using different method.
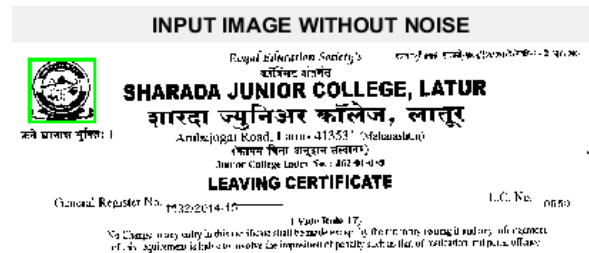
for n=1:size
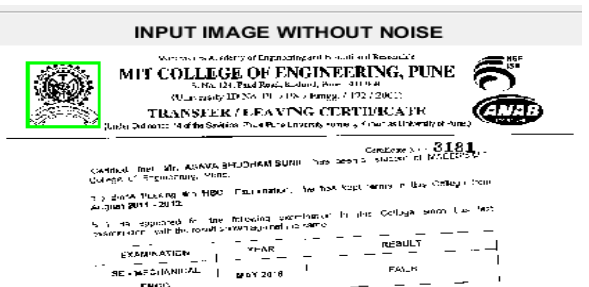apply function
end

Fig .9 Select logo from original image

Fig. 10 Select logo from original image

Published by :

http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 05, May - 2017

Step 5 Extract the logo from Document
for n=1:n
   [r,c] = find(L==n);
   n1=imagen(min(r):max(r),min(c):max(c));
   imshow(~n1);
   pause(0.5)
end

Fig. 11 Extracted logo from original image

Fig (1,2,3,4) Is the process of extracting Logo from
Transfer Certificate.
Step 6 Input object image and scan image.
Step 7 Convert RGB to gray
Step 8   Apply detectMSERFeatures function extract region.
Step 9 Extract feature by using extract Features function
Step 10match feature by using matchFeatures function.
Step 11Show match feature in following figure
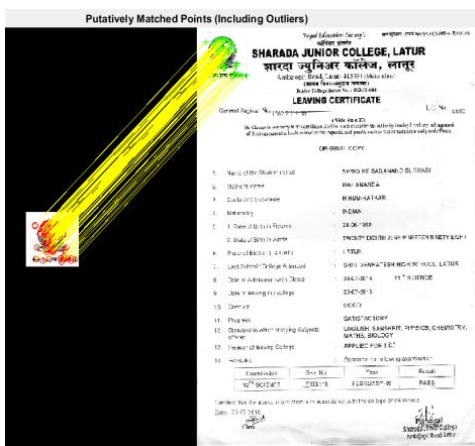

Fig. 12 Match feature including outliers

Fig. 13 Match feature including Inliers

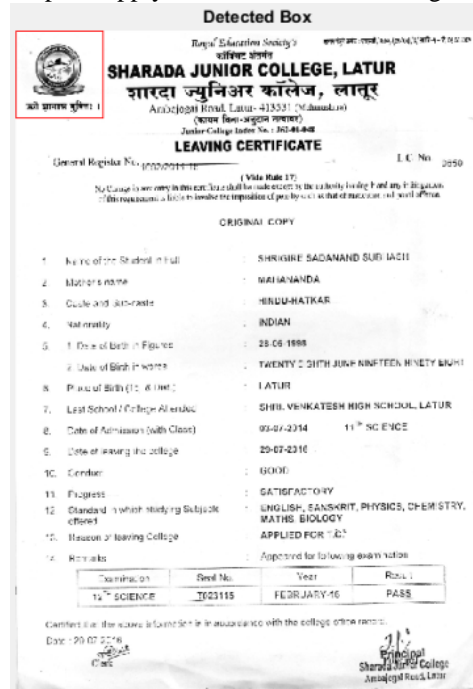Step 12 Apply line function to select logo from image

Fig .14 Select logo from image

F.   There are four major issues in Document Verification

Logo detection: to extract the logo from different transfer certificate images of different organization.
Feature Extraction: to extract the feature from extracted logo.
Feature Matching: to match the feature of logo with original images.
Document verification: to verify the originality of document.
Application Area:
Proposed module is developed to verify the document

VI.   RESULT AND DISCUSSION

This paper pass RGB image as input of TC then convert the RGB image to gray scale image then select the particular logo from image extract the logo from TC with full accuracy. This paper pass different images of different institution with different size like 3500 x 2500 to the module and this module return only extracted logo from large images. And the logo size is 400x400 with full accuracy. This paper applies SURF method to extract and match the logo. Spatial density method is also used to extract the logo from images. But surf method is better than this method.

VII.   CONCLUSION

The purpose of this paper is to extract the logo from different TCs, match the feature of logo with original TC or any other document if it is match then it display the result weather is it original or duplicate or if it is match then display the name of college. Segmentation based on Gaussian filter, edge detection, content-based image retrieval, thresholding and histogram-based methods. All these techniques are backbone for logo segmentation also planned to design a novel approach for region detection and segmentation for the document images. This paper presents the module to extract the logo automatically from different document of different institution for document verification. Purpose is that it could find the particular document is belonging to which institute or university. Proposed work is used to apply different parameters to check originality of document like logo on document, reference no, Name of candidate, issue date of document, stamp on document for the verification.

REFERENCES

[1] Meera Sambhaji Sawalkar1, Mrunalinee Patole2, "Logo Detection and Recognition from the Images as Well as Videos" in Proceedings of International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Impact Factor (2012): 3.358

[2] S.Balan, Dr.P.Ponmuthuramalingam "Automatic Web Page Logo Detection" in Proceedings of International conference on Signal Processing, Communication, Power and Embedded System (SCOPES)-2016

[3] Remya Ramachandran, Andrews Jose "Logo Matching And Recognition System Using Surf" in Proceedings of International Journal of Research in Computer and Communication Technology, Vol 3, Issue 9, September – 2014

[4] 1Prof. Mrunalinee Patole, 2 Meera Sambhaji Sawalkar, "Improved approach for Logo Detection and Recognition" in Proceedings of International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) Volume 3, Issue 6, November-December 2014

[5] B.Sai Prasanthi1, D. Rajesh Setty2 "Indian Paper Currency Authentication System using Image processing" in Proceedings of International Journal of Scientific Research Engineering & Technology (IJSRET), ISSN 2278 – 0882   Volume 4, Issue 9, September 2015

[6] 1Mustafa S. Kadhm, Lim Shu Yun 2, "Propose A Simple and Practical Vehicle Logo Detection and Extraction Framework" in Proceedings of International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) Volume 4, Issue 3, May-June 2015

[7] Mrs.Prajkata Musale, Prof.Mr. Deepak Gupta, "Improve speed of Logo Detection and Recognition from the images Using SURF" in Proceedings of International Journal of Advancements in Research & Technology, Volume 4, Issue 12, December -2015   1 ISSN 2278-7763

[8] Fanman Meng, Hongliang Li, "From Logo to Object Segmentation" in Proceedings of IEEE Transactions On Multimedia, Vol. 15, No. 8, December 2013

[9] Engr.Farman Ali1, Engr.Sajid Ullah Khan2, Engr.Muhammad Zarrar Mahmudi3, Rahmat Ullah4 "A Comparison of FAST, SURF, Eigen, Harris, and MSER Features" International Journal of Computer Engineering and Information Technology VOL. 8, NO. 6, June 2016, 100–105 Available online at: www.ijceit.org E-ISSN 2412-8856 (Online)

[10] Remya Ramachandran, Andrews Jose "Logo Matching And Recognition System Using Surf" International Journal of Research in Computer and Communication Technology, Vol 3, Issue 9, September - 2014 ISSN (Online) 2278- 5841