Text and Texture Extraction for Various Images using Combinational Methods

Kusum Jonwal¹ Jyoti Mahodia²
¹,²M.E. Student
¹,²G.S. Inst of Technology and Science Indore (MP)-452003, India

Abstract—Text data contains useful information for automatic explanation, indexing, and structuring of images. The typical low-level features that are extracted in images and video include measures of color, texture, or shape. Extraction of this information involves detection, localization, tracking, extraction, enhancement, and recognition of the text from a given image. The objective of this paper is to compare three basic approaches of text and texture extraction in natural images namely; edge-based, connected-component based and text based method. The edge based and connected component based method shows how to extract text from an image whereas the text extraction method shows how to extract texture information from an image here creating a GLCM, and then extracting statistical measures from this matrix. Further, the image entropy and energies are also computed in order to correlate the skin symptoms to the skin texture images.

Key words: Text, Texture, TIE, GLCM

I. INTRODUCTION

The Text in images contains important and useful information. Its content can be in the form of objects, color, texture, shape as well as the relationships between them. The text data can be embedded in an image or video in different font styles, sizes, orientations, colors, and against a complex background. Text extraction is the ability to extract (pull out) text from a document or image. The text has some common characteristics in terms of frequency and orientation information, and also spatial cohesion. Spatial cohesion refers to the fact that text characters of the same string appear close to each other and are of similar height, orientation and spacing.

A. Text Information Extraction (TIE)

This is a system which receives an input in the form of an image. The images can be in color or gray scale, compressed or uncompressed, and the text in the images may move or may not. The problem arises due to TIE system can be divided into the following sub-problems: (i) detection (ii) localization (iii) tracking (iv) extraction and enhancement (v) recognition (OCR). Text detection refers to the determination of the presence of text in a given sequence of images. Text localization is the process of determining the location of text in the image. Text tracking is performed to reduce the processing time for text localization. The text need to be segmented from the background to facilitate its recognition. Which means, the extracted text image has to be converted into a binary image and enhanced before it is fed into an OCR engine. Text extraction is the stage where the text components are segmented from the background. Text extracted text images can be transformed into plain text using OCR technology.

B. Texture Information Extraction (TIE)

Texture analysis is an important tool in analyzing the image of textural nature. The skin texture is the appearance of the skin smooth surface. The features like amount of collagen and hormones, diet and hydration can be analyzed by textural information. The texture appearance is changing with image recording parameters, that are camera, illumination and direction of view, a problem common to any real surface. In the digital image processing, several methods have been developed to classify images and define statistical distances among them, with the aim to decide whether, in a set of many images, there exist some which are close to any arbitrary image previously encountered. The texture discrimination can be obtained by choosing a set of attributes, the texture features, which account for the spatial organization of the image.

II. RELATED WORKS

A. Text Extraction Techniques

To implement, test, and compare and contrast. Approaches for text region extraction in images, and these are used to find how the algorithms perform under variations of lighting, orientation, and scale transformations of the text.

1) Edge based methods focus on the high contrast between the text and the background. The edges of the text boundary are identified and merged, and then used to filter out the non-text region. Once the region is recognized, spatial cohesion features are applied in order to discard false positives. where the spatial cohesion Spatial cohesion refers to the fact that text characters of the same string appear close to each other and are of similar height, orientation and spacing.

2) Connected Component based methods uses a bottom-up approach by grouping small components into successively larger components until all regions are identified in the image.

3) Text based extraction method: text based method is a feature based algorithm which involves the construction of gray level co-occurrence matrix (GLCM). These matrix are used to calculate the features like contrast, homogeneity, dissimilarity and which are the results for feature extraction in texture based method. The text regions have special texture. Though these methods are comparatively less sensitive to background colors, they may not differentiate the texts from the text-like backgrounds.

III. METHODOLOGY

In this paper we are discussed about edge based connected component based method, but by using this method we unable to extract text properly. Cause that we are using
connected component based method. In which we follow some steps.
1) Convert the input image to YUV color space. The luminance (Y) value is used for further processing. The output is a gray image.
2) Convert the gray image to an edge image.
3) Compute the horizontal and vertical projection profiles of candidate text regions using a histogram with an appropriate threshold value.
4) Use geometric properties of text such as width to height ratio of characters to eliminate possible non-text regions.
5) Binarize the edge image enhancing only the text regions against a plain black background.

The vertical projection profile shows the sum of pixels present in each column of the intensity or the sharpened image. And the horizontal projection profile shows the sum of pixels present in each row of the intensity image. These projection profiles are essentially histograms where each bin is a count of the total number of pixels present in each row or column. Candidate text regions are segmented based on adaptive threshold values, $T_1$ and $T_2$, calculated for the vertical and horizontal projections respectively. Only regions that fall within the threshold limits are considered as candidates for text. The value of threshold $T_1$ is selected to eliminate possible non-text regions such as doors, window edges etc. that have a strong vertical orientation. Similarly, the value of threshold $T_2$ is selected to eliminate regions which might be non-text or long edges in the horizontal orientation.

\[
T_x = \frac{\text{Mean (Horizontal projection profile)}}{20}
\]

\[
T_y = \text{Mean (Vertical projection profile)} + \max(\text{Vertical projection profile})
\]

![Fig. 1: (a) Original Image (b) Gray scale Image (c) Result Image](image)

The performance of each technique has been evaluated based on its precision and recall rates obtained.

Texture is examining by the statistical method that considers the spatial relationship of pixels is the gray level co-occurrence matrix (GLCM). The GLCM function characterize the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. The thesis work is divided into following steps:

1) Image Acquisition
2) Conversion to Gray Scale Image
3) Image Enhancement using Histogram Equalization
4) Histogram Computation of the enhanced image
5) Computation of GLCM Matrix of skin texture image
6) Computation of Contrast
7) Computation of Entropy
8) Computation of Energy
9) Computation of Homogeneity
10) Correlation with Skin Symptoms

![Fig. 4: Methodology of human skin texture analysis using image processing techniques](image)
IV. HISTOGRAM EQUIVALENT IMAGE

![Images showing original and histogram equalized images]

Fig. 5: Images with their histogram equivalent

V. GLCM EXTRACTION

This is a statistical method of examining texture that shows the spatial relationship between pixels which can be also known as gray-level spatial dependence matrix. The GLCM functions characterize the texture of an image by calculating how often pairs of pixel with specific values and specified spatial relationship occur in an image. Creating a GLCM and then extracting statistical measure from this matrix. The spatial relationship is defined as the pixel of interest and the pixel to its immediate right horizontally adjacent. To create a GLCM, use the graycomatrix function. The graycomatrix function creates a gray-level co-occurrence matrix (GLCM) by calculating how often a pixel with the intensity (gray-level) value \( i \) occurs in a specific spatial relationship to a pixel with the value \( j \). By default, the spatial relationship is defined as the pixel of interest and the pixel to its immediate right (horizontally adjacent). After you create the GLCMs, image contrast, energy, correlation and homogeneity can be computed as:

- **Contrast**: This is statistic measures the spatial frequency of an image and is difference moment of GLCM. It is the difference between the highest and the lowest values of a contiguous set of pixels. Contrast should zero for constant image.
- **Correlation**: It measures the joint probability occurrence of the specified pixel pairs. Correlation is 1 or -1 for perfectly positive and perfectly negative correlated image. The correlation feature is a measure of gray tone linear dependencies in the image.
- **Energy**: It provides the sum of squared elements in the GLCM also known as uniformity or the angular second moment. Energy is 1 for a constant image. It shows how gray levels are distributed.
- **Homogeneity**: It measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Homogeneity is 1 for a diagonal GLCM.
- **Entropy**: This statistic measures the disorder or complexity of an image. The entropy is large when the image is not texturally uniform and many GLCM elements have very small values.

\[
H = - \sum_{i=0}^{L-1} p_i \ln p_i,
\]

Here \( L \) denotes the number of gray level. \( p_i \) equals the ratio between the number of pixels whose gray value equals \( i \) and the total pixel number contained in an image.

A. Generating Skin Map:

The region to generate skin map is to highlight patches of “skin” like pixels. This can be used in face detection and gesture recognition. Here Out - contains the skin map overlayed onto the image with skin pixels marked in blue color. Bin - contains the binary skin map, with skin pixels as ‘1’. For texture characterization, we consider a set of features derived from GLCM matrix: contrast (C), homogeneity (H), mean (M), energy (N). Images are obtained from Dermnet Skin disease atlas. Dermnet is the largest independent photo dermatology source. Dermnet provides information on a wide variety of skin conditions. The proposed algorithm produce a skin map of a given image and highlights patches of skin like pixels. A skin map overlayed onto the image with skin pixels marked in blue color is generated by using the GLCM matrix. Following figures shows the output of the algorithm. Once skin pixels are extracted, than it is easier to analyze the skin diseases.

![Fig. 6: Generating Skin Map]
VI. CONCLUSION

The main focus of this paper is on extraction of text and texture from an image and that can be use in security purpose and diagnose the skin diseases. An accurate text region extraction algorithm based on three methods with gray-information is presented. Here the connected component based method is used for text extraction and text based method is used for texture extraction. Various skin diseases can be analyzed based on the combination of feature vector set of contrast, correlation, energy and homogeneity. From the experimental results discussed above, we infer that the multi-class classification can serve as an effective tool in identifying skin diseases. The future work will be based on developing algorithms for reduce the complexity of text extraction and to identify various other skin diseases, to improve the accuracy and reduction of percentage error.

Results for contrast, correlation, energy and homogeneity are summarized in below given table.

<table>
<thead>
<tr>
<th>Image</th>
<th>Contrast</th>
<th>Correlation</th>
<th>Energy</th>
<th>Homogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.8926</td>
<td>0.7843</td>
<td>0.0314</td>
<td>0.6372</td>
</tr>
<tr>
<td>B</td>
<td>0.2542</td>
<td>0.4126</td>
<td>0.4498</td>
<td>0.8892</td>
</tr>
<tr>
<td>C</td>
<td>0.8882</td>
<td>0.8851</td>
<td>0.0517</td>
<td>0.7498</td>
</tr>
<tr>
<td>D</td>
<td>0.4269</td>
<td>0.9525</td>
<td>0.0747</td>
<td>0.8665</td>
</tr>
<tr>
<td>E</td>
<td>0.3429</td>
<td>0.9670</td>
<td>0.1043</td>
<td>0.9025</td>
</tr>
</tbody>
</table>

Table 1: Results for contrast, correlation, energy and homogeneity

REFERENCES

[6] Lei Huang1,2, Tian Xia1, Yongdong Zhang1, Shouxun Lin1, ‘HUMAN SKIN DETECTION IN IMAGES BYMSER ANALYSIS”, IEEE-2011.