AN APPORACH FOR SCRIPT IDENTIFICATION IN PRINTED TRILINGUAL DOCUMENTS USING TEXTURAL FEATURES

Mahesha D M and Gopalan N P

Department of Computer science, Centre for Research and Development, PRIST University Thanjavur, Tamilnadu, India.
Department of Computer Applications NIT, Trichirappalli, Tamilnadu, India.

ABSTRACT

In this work, we review the outcome of texture features for script classification. Rectangular White Space analysis algorithm is used to analyze and identify heterogeneous layouts of document images. The texture features, namely the color texture moments, Local binary pattern (LBP) and responses of Gabor, LM-filter, S-filter, R-filter are extracted, and combinations of these are considered in the classification. In this work, a probabilistic neural network and Nearest Neighbor are used for classification. To corrabate the adequacy of the proposed strategy, an experiment was operated on our own data set. To study the effect of classification accuracy, we vary the database sizes and the results show that the combination of multiple features vastly improves the performance.

KEYWORDS

Segmentation, Section finding, Section Merge, Feature Extraction, Classification.

1. INTRODUCTION

Anything which conveys information is known as a document. Generally, a document is a knowledge container. Most of the times we acquire knowledge from documents such as Newspapers, Textbooks, Scientific journals, Magazines, Technical reports, Office files, Postal letters, Bank cheques, Application forms etc. (Tang et al., [1]). To understand the huge information, an extensive amount of manual processing is required and such a manual processing is very much time consuming. To overcome this difficulty, it is essential to automate the manual process which needs efficient algorithms. This automation process is considered as document image processing (DIP). In general, the document image processing is divided into text processing and graphics processing. Text processing is further divided into character recognition and page layout analysis. Graphics processing is further divided into line processing and region processing as shown in Figure 1.

1.1 Stages in Document Image Processing

The document image processing involves three basic steps at conceptual levels, which are document image analysis, document image recognition and document image understanding. Within these three levels, there are several other interacting modules such as image acquisition, binarization, block segmentation, block classification, logical block grouping, character and word recognition, picture processing and analysis, graphic analysis, picture understanding, text understanding and graphics understanding. The interactions between these processes and data flow between levels are shown in Figure 2.

1.1.1   Document Image Analysis(DIA)

DIA is a process of recovering syntactic and semantic information from images of documents, prominently from scanned versions of paper documents. Two unique tasks in document image analysis. The first is identification of basic components of the document, the so-called document objects. The second one is consisting of the identification of role and meaning of the document objects in order to have an interpretation of the whole original document. The structural analysis, on the other hand involves usage of layout clues to identify headlines, locate different lines, etc. In general, image analysis involves the extraction and use of attributes and structure relationships in the document in order to label its components within contextual rules dictated by the document class. Analysis of printed documents obviously involves skew angle estimation and correction which is a very challenging task.

1.1.2 Document Image Recognition (DIR)

DIR is valuable method in office robotization and advanced library applications, is to locate the most comparable layout for any information archive picture in a prestored format record picture. These days a lot of existing paper records are changed to computerized archive pictures through scanners and cameras. However, the next step is to analyze a document and segregate text blocks, graphic block, picture block, etc, so as to facilitate labeling of the blocks. This process of labeling the blocks is said to be document image recognition or identification.

1.1.3 Document Image Understanding

Document image understanding is a component which extracts the logical relationships between the respective blocks of a document. Logical document structure is a hierarchical representation of semantics of the given document. The same logical document structure is formatted in varieties of physical layouts by changing the variables such as number of pages and font sizes, spacing between paragraphs and between sections, number of columns, etc. In all these layouts, the semantics of the document remains unaltered. Document's semantic structure can be determines by Logical structure analysis and also it provides data appropriate for information retrieval.
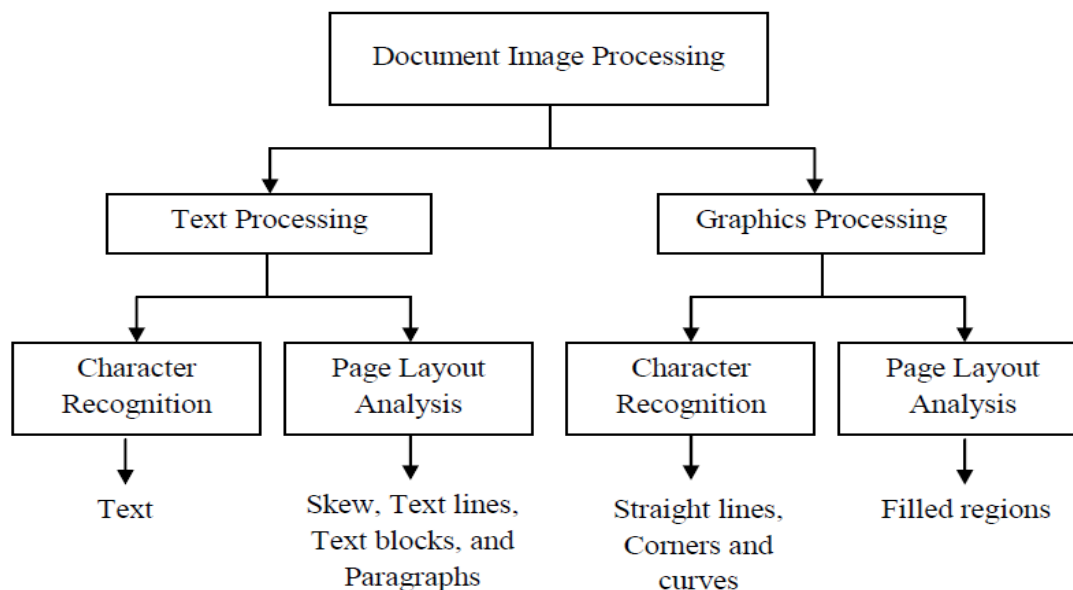
Figure 1. Hierarchy of document image processing with subcategories

1.2 Script Recognition

The OCR innovation for Indian archives is in developing stage and the vast majority of these Indian OCR frameworks can read the reports written in just a solitary script. According to the Indian constitution, each state Government needs to create an official archive containing a national dialect (Hindi), official dialect (English) and state dialect (or local dialect). As per the three-dialect strategy embraced by the greater part of the Indian expresses, the records delivered in an Indian state Karnataka, are made out of writings in the local dialect Kannada, national dialect Hindi and the overall regularly utilized dialect English. What's more, greater part of the reports found in a large portion of the private and Government segments of Indian states, are Trilingual sort (an archive having content in three dialects). Thus, there is a developing interest to consequently prepare these Trilingual records in each state in India, including Karnataka.

The monolingual OCR frameworks won't process such multi-script archives without human contribution for outlining distinctive script zones of multi-lingual pages before enacting the script particular OCR motor. The requirement for such manual contribution can bring about more noteworthy cost and vitally postpones the general picture to-content change. Consequently, a programmed sending is required for the approaching record pictures to handover this to the specific OCR motor contingent upon the information of the characteristic scripts. In perspective of this, ID of script and/or dialect is one of the basic assignments for multi-script record handling. A script recognizer, thusly, streamlines the errand of OCR by upgrading the precision of acknowledgment and lessening the computational intricacy.

Script Recognition approaches can be broadly classified into two categories, namely, local and global approaches. The neighborhood approaches(Pal and Chaudhury[2], Pal et al [3]) dissect a rundown of associated parts (Line, word, scorch) in the archive pictures, to distinguish the script(or class of script). Interestingly, worldwide methodologies (Joshi [4]) utilize an examination of districts (square of content) containing atleast two lines (or words)without better division. As a rule, worldwide methodologies function admirably taking into account surface estimation, yet this depends vigorously on a uniform piece of content (Buschet al [5]), and broad preprocessing (to make the content square uniform) is required to gauge the surface. Despite the fact that neighborhood approaches depend on the precision of character division or associated part investigation, it could function admirably on the reports regardless of their quality or consistency in the piece of content.

In the literature, many works have been reported for script recognition at the document, line and word levels, using local approaches. In this context, researchers have made a number of attempts to discriminate the Han and Latin script (Spitz [6], Lu and Tan [7]) at the document level and exploited many Indian scripts at line level and word level (Pal and Chaudhury[8], Pal and Chaudhury [9], Padma and Nagabhushan[10], Dhandra et al [11]). However, all the techniques reported in the literature are script dependent. Since this research is intended to develop an classification system for kannada document images, Script Recognition, to discriminate the kannada from English scripts in bilingual document images is becoming important. In this connection, few local approaches are reported in the literature, such as spatial spread analysis (Dhanya et al [12]), Aspect Ratio (Tan et al [13]), Structural features (Pal and Chaudhury [14]), and Water Reservoirs (Pal et al [15]).However, all the above mentioned techniques produce a low discrimination rate due to its incapability in exploration of thescripts. Global approaches (Pati et al [17], Pati and Ramakrishnan [18]). S Chaudhury et al., [19] has proposed a method for identification of Indian languages by combining Gabor filter based technique and direction distance histogram classifier considering Hindi, English, Malayalam, Bengali, Telugu and Urdu.
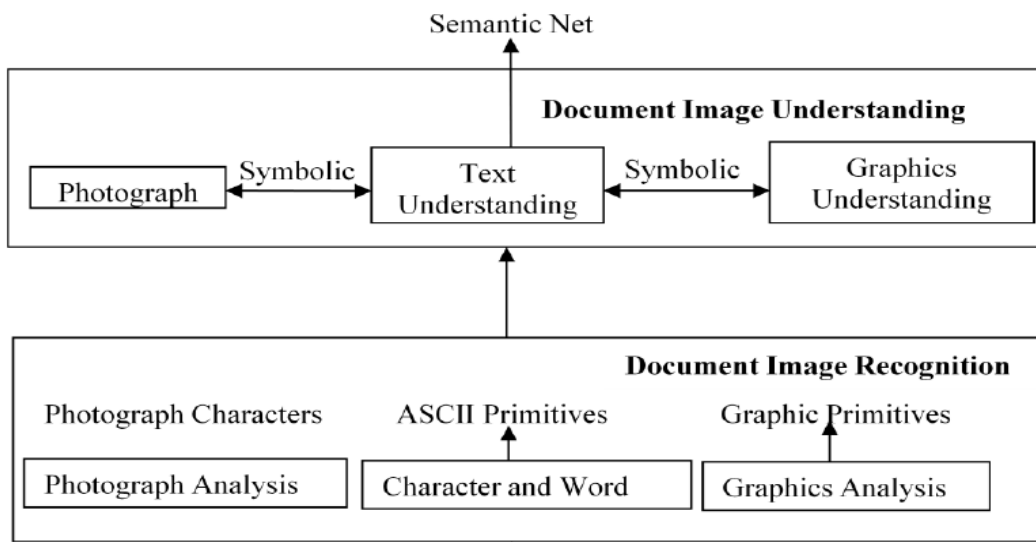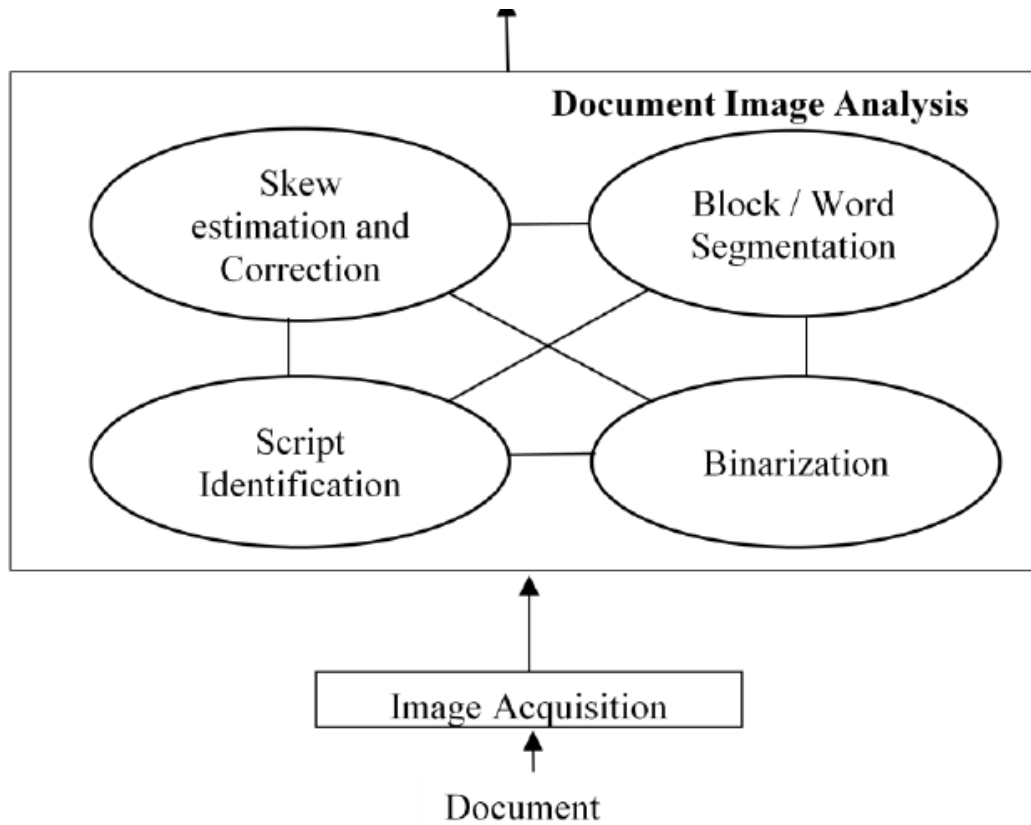
Figure 2. Steps involved in document image processing

G D Joshi et al., [20] built up a script recognizable proof method for 10 Indian scripts utilizing an arrangement of components extricated from log Gabor channels. Dhanya et al., [21] have utilized Linear Support Vector Machine (LSVM), K-Nearest Neighbor (K-NN) and Neural Network (NN) classifiers on Gabor-based and zoning elements to characterize Tamil and English scripts. Hiremath [22] proposed a methodology for script ID of South Indian scripts utilizing histogram highlights. S R Kunte and S Samuel [23] have utilized a neural methodology as a part of on-line script acknowledgment for Telugu dialect utilizing wavelet highlights. Peeta et al., [24] have exhibited a method utilizing Gabor channels for script recognizable proof of Indian bilingual archives.

2. SEGMENTATION

The preprocessing of the document images includes the process of Noise removal and Binarization. Noise could occur in the document images due to many sources such as aging, photocopying etc. and the application of filters reduces noises in these images. Here noise has been suppressed in the document image by using a median filter, since median filters smear the character image strokes. For this median filter, a 3*3 mask has been chosen and it is applied over the image, which replaces nine pixels by the intensity of the center pixel over this mask.As a

result of pulling the median filter output to the gray level of the center pixel, the shapes of the character strokes can be preserved. Binarization has been applied after noise removal. Binarization is a technique by which the color and gray scale images are converted into binary images. The most widely recognized strategy is to choose an appropriate threshold for the picture and change over all. The intensity values above the threshold value representing as 'white' and below the threshold as 'black' value.

Since all the books and magazines use white spaces as a separator within and between the texts, the observation of small white spaces becomes mandatory to identify the text area. Therefore, in Section Finding, white spaces are used as delimiters and observed for analysis. Variable length white spaces exist inside the text in both the directions, apart from the white spaces surrounding the textual zones. Due to the existence of non-uniform, small white gaps in the image apart from the column separators, a careful analysis is required to observe and record the white spaces. As a result, the width of the image has been divided into 'n' equal sections. Since connected component analysis has been eliminated, a single horizontal scan has been performed over the image to grab the white spaces. After an entire horizontal scan of an image, all the sections which appear as white spaces are reported and their positions with the corresponding row number have been recorded as a result of this procedure. It is hard to process various white space section numbers to identify the layout gaps if the merging procedure has been avoided. Once all the white space section numbers based on their row number have been indicated, the merging of adjacent sections in both the directions is required to form horizontal and vertical white space rectangles which are done through theSection Merging phase.

The Section Merging phase consists of two processes: Horizontal Section Merging and Vertical Section Merging. Initially, horizontal section merging accepts all the white space section numbers with their corresponding row numbers as the input and produces a series of within-line or row-wise white space clusters as output(i.e.), subsequent white space sections in each row gets merged together to produce a series of row-wise white space sections. Since all the white spaces (section-wise) are identified and merged properly, the chance of getting under-segmentation has been completely eliminated.The rectangular analysis phase consists of Cropping and the Rectangular formation process. After the identification of horizontal and vertical white space rectangles, finding the areas which are uncovered by the white space rectangles could yield the layout.The Cropping procedure acts over the white space rectangles in both the directions by accepting the horizontal edges of each Horizontal White Space Rectangle (HWSR) and the vertical edge of each Vertical White Space Rectangle (VWSR).

Once the horizontal and vertical edges are cropped, the areas uncovered by the white spaces could be easily extracted through rectangular formation procedure.Once the content blocks have been identified, the next step attempts to separate the textual blocks from the images and pictures, since textual blocks are required for further processing. Once the homogeneous regions are obtained, each region gets passed into the text image analyzer to identify the text component. Two statistical properties called as Black Run Length (BRL) and White Black Transition Count (WBTC), which spans in the horizontal direction of the image have been used here to identify the textual blocks. Black run length corresponds to the ratio of the total number of black pixels in a row to the total transition (black-white disposition) count in that row. The White Black Transition count corresponds to the ratio of the total number of transitions in a row to the total number of

pixels in that row. It is concluded that if the mean black run length appears to be more, and the Mean white black transition count of all the rows appears to be lesser than the threshold, it is concluded as image and not as a text.
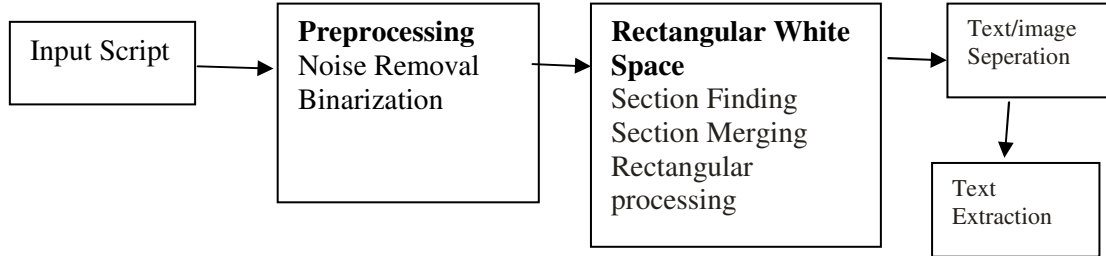

Figure 3. Steps involved in text extraction

ವೆಸ್ಟ್ ಇಂಡೀಸ್ ಕ್ರಿಕೆಟ್ ಮಂಡಳಿ ಜತೆ ಆಟಗಾರರ ತಿಕ್ಕಾಟ ಮುಂದುವರೆದಿರುವುದರಿಂದ ಟೆಸ್ಟ್ ಕ್ರಿಕೆಟ್ ಆಗಿ ಸಂಪಾದನೆ ಮಾಡುವುದು ಸಾಕಾಗುತ್ತಿಲ್ಲ. ಕಡಿಮೆ ಸಂಭಾವನೆ ಪಡೆದು ಟೆಸ್ಟ್ ಆಡುವುದಕ್ಕಿಂತ ಟಿ20 ಅಂತಾರಾಷ್ಟ್ರೀಯ ಪಂದ್ಯ ಹಾಗೂ ಕೆರಿಬಿಯನ್ ಲೀಗ್ ಆಡುವುದು ಉತ್ತಮ ಎಂದು ಸ್ಯಾಮುಯೆಲ್ಸ್ ಹೇಳಿಕೊಂಡಿದ್ದಾರೆ.[12 ಸಾವಿರ ರನ್ ಕ್ಲಬ್ ಸೇರಿದ ವಿರಾಟ್ ಕೊಹ್ಲಿ]

Figure 4. Input Image with text

ವೆಸ್ಟ್ ಇಂಡೀಸ್ ಕ್ರಿಕೆಟ್ ಮಂಡಳಿ ಜತೆ ಆಟಗಾರರ ತಿಕ್ಕಾಟ ಮುಂದುವರೆದಿರುವುದರಿಂದ ಟೆಸ್ಟ್ ಕ್ರಿಕೆಟ್ ಆಗಿ
ಸಂಪಾದನೆ ಮಾಡುವುದು ಸಾಕಾಗುತ್ತಿಲ್ಲ. ಕಡಿಮೆ ಸಂಭಾವನೆ ಪಡೆದು ಟೆಸ್ಟ್ ಆಡುವುದಕ್ಕಿಂತ ಟ20
ಅಂತಾರಾಷ್ಟ್ರೀಯ ಪಂದ್ಯ ಹಾಗೂ ಕರಿಬಿಯನ್ ಲೀಗ್ ಆಡುವುದು ಉತ್ತಮ ಎಂದು ಸ್ಯಾಮುಯಲ್ಸ್
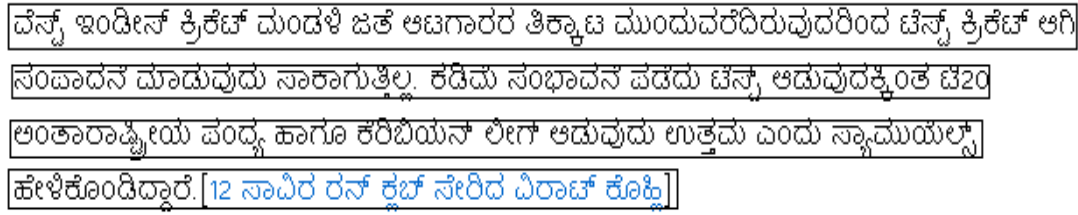ಹೇಳಿಕೊಂಡಿದ್ದಾರೆ. [12 ಸಾವಿರ ರನ್ ಕ್ಲಬ್ ಸೇರಿದ ವಿರಾಟ್ ಕೊಹ್ಲಿ]

Figure 5. Input Image with elimination text

3. FEATURE EXTRACTION

As our significance is to study surface elements for grouping, for a portioned picture we separate Color Texture Moments, Local twofold example, LM channel reactions, Schmid channel reactions, Maximum channel reactions and Gabor reactions for surface examination. The streaming subsection gives a prologue to all the above surface elements.

3.1 Gabor Filter Responses

The Texture investigation utilizing channels in view of Gabor capacities falls into the classification of recurrence based methodologies. These methodologies depend on the reason that surface is a picture design containing a dull structure that can be viably portrayed in a recurrence area, for example, the Fourier space. One of the difficulties, notwithstanding, of such a methodology is managing the tradeoff between the joint vulnerability in the space and recurrence areas. An attractive mathematical property of Gabor functions is that they minimize the joint uncertainty in space and frequency. They achieve the optimal tradeoff between localizing the analysis in the spatial and frequency domains. Using Gabor filters to analyze texture appeals from a psycho-visual perspective as well. The texture analysis is accomplished by applying a bank of scale and orientation selective Gabor filters to an image [25]. These filters are constructed as follows. A two-dimensional Gabor function $g(x; y)$ and its Fourier transform $G(u; v)$ can be written as:

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi iWx\right]$$

and

$$G(u,v) = \exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\}$$

Where $i = \sqrt{-1}$, $\sigma_u = \frac{1}{2\pi\sigma_x}$ and $\sigma_v = \frac{1}{2\pi\sigma_y}$ control the tradeoff between spatial and frequency resolution, and W controls the modulation. A class of self-similar functions referred to as Gabor wavelets is now considered. Let $g(x, y)$ be the mother wavelet. A filter dictionary can be obtained by appropriate dilations and translations of $g(x, y)$ through the generating function.

$$g_{rs}(x,y) = a^{-s}g(x',y'), \; a > 1, \; s \, \varepsilon \, 0,1..S - 1, \; r \varepsilon 1,2,..R$$

$$x' = a^{-s}(x\cos\theta + y\sin\theta) \text{ and } y' = a^{-s}(-x\sin\theta + y\cos\theta)$$

- where $\theta = (r-1)\pi/R$. The indices r and s indicate the orientation and scale of the filter respectively. R is the total number of orientations and S is the total number of scales in the filter bank.

3.2 Maximum Response (MR) Filter Responses

We depict the surface by convolving the pictures with MR channel bank presented by Varma and Zisserman[26]. The MR sets contain both isotropic channels and additionally anisotropic channels at various introductions they produce great components for a wide range of surfaces. Moreover, not at all like customary rotationally invariant channels, the MR sets are likewise ready to record the point of most extreme reaction. This empowers to register higher request co-event measurements on introduction and such insights may demonstrate helpful in separating surfaces which give off an impression of being fundamentally the same as. The MR channel banks create more critical textons not just as a result of enhanced grouping in a lower dimensional space additionally on the grounds that pivoted components are accurately mapped to the same texton.

3.3 Leung-Malik (LM) and Schmid Filter Responses

The LM set is a multi scale, multi orientation filter bank with 48 filters. It consists of first and second derivatives of Gaussians at 6 orientations and 3 scales making a total of 36; 8 Laplacian of Gaussian (LOG) filters; and 4 Gaussians. In [27] two versions of the LM filter bank are considered. In LM Small (LMS), the filters occur at basic scales $\sigma = \{1, \sqrt{2}, 2, 2\sqrt{2}\}$. The first and second derivative filters occur at the first three scales with an elongation factor of 3 (i.e. $\sigma_x = \sigma$ and $\sigma_y = 3\sigma_x$). The Gaussians occur at the four basic scales while the 8 LOG filters occur at σ and 3σ. For LM Large (LML), the filters occur at the basic scales $\sigma = \{\sqrt{2}, 2, 2\sqrt{2}, 4\}$. A Schmid response [14] is rotational invariant but the invariance is achieved in a different manner and texton clustering is always in a higher dimensional space. The Schmid set consists of 13 rotationally invariant filters of the form.

$$F(r,\sigma,\tau) = F_0(\sigma,\tau) + \cos(\frac{\pi\tau r}{\sigma})e^{-\frac{r^2}{2\sigma^2}}$$

3.4 Color Texture Moments (CTM)

Yu et al., in their work [28] developed a color and texture combined moments for image retrieval. In their work the image was converted onto different color space viz., RGB, HSV, YUV, and (SVcosH, SVsinH, V). Local Fourier Transform (LFT) was performed on all the channels of image with eight different templates and for each resulted channel image first two moments were calculated obtaining 48 features for a given single image. Experimentally they found out that

(SVcosH, SVsinH, V) color space is better than other color space and these features were termed as color texture moments. In our work we have used this color texture moments as one of the features.

3.5 Local Binary Pattern

Ojala et al. [29] proposed to utilize the Local Binary Pattern (LBP) histogram for revolution invariant surface grouping. LBP is a straightforward however proficient administrator to depict nearby picture designs. It is consolidated factual and organized technique. The LBP administrator considers a 3x3 neighborhood of a pixel, then these 8 outskirt pixels will be supplanted either by 1, in the event that they are bigger than or equivalent to the focal pixel or by 0 generally. At long last, the focal pixel will be supplanted with a summation of the double weights of outskirt pixels in the LBP picture and the 3x3 window slides to the following pixel.

It is possible to develop the basic LBP into various neighborhood sizes and distances.

$$LBP_{P,R} = \sum_{P=0}^{P-1} s(g_P - g_c)2^P \qquad (1)$$

Where $s(.)$ is the sign function:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \qquad (2)$$

g_p and g_c are grey levels of border pixels and central pixel respectively, P is the number of pixels in the neighborhood and R is the radius of the neighborhood. Suppose the coordinates of g_c are (0, 0), then the coordinates of g_p are given by (-R sin (2πp/P), R cos (2πp/P)).

In this case, if we set ($P = 8$; $R = 1$), we obtain the basic LBP (1) Luminance changing cannot affect signed differences g_p-g_c, hence LBP is grey level shift invariant.

6	5	2
7	6	1
13	28	11

1	0	0
1		0
1	1	1

1	2	4
128		8
64	32	16

	241	

Image   ThresholdWeightsLBP=1+16+32+

64+128=241
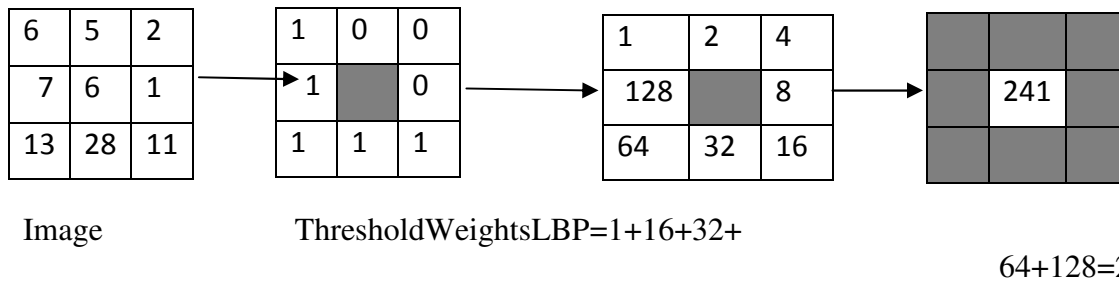
Fig 6: Computing the basic 3x3 LBP.

Suppose the texture image is NxM. After identifying the LBP pattern of each pixel (i, j), the whole texture image is represented by building a histogram:

$$H(k) = \sum_{i=1}^{N} \sum_{j=1}^{M} f(LBP_{P,R}(i,j),k), k \in [0,K] \quad (3)$$

$$f(x,y) = \begin{cases} 1, & x = y \\ 0, & otherwise \end{cases} \quad (4)$$

where K is the maximal LBP pattern value. The U value of an LBP pattern is defined as the number of spatial transitions (bitwise0/1 changes) in that pattern

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{P=1}^{P-1} |s(g_P - g_c) - s(g_{P-1} - g_c)| \quad (5)$$

For example LBP pattern 00000000 has a U value of 0 and 01000000 of 2.The uniform LBP pattern refers to the uniform appearance pattern which has limited transition or discontinuities (U≤2) in the circular binary presentation. It verified hat only ''uniform'' patterns are fundamental patterns of local image texture.In practice, the mapping from $LBP_{P,R}$ to$LBP^{u2}_{P,R}$ (superscript "$u2$" means uniform patterns with $U{\leq}2$),which has $P*(P-1)+3$ distinct output values, is implemented with a lookup table of 2^Pelements.

To achieve rotation invariance, a locally rotation invariant pattern could be defined as:

$$LBP_{P,R}^{riu^2} = \begin{cases} \sum_{P=0}^{P-1} s(g_P - g_c) & if \ U(LBP_{P,R}) \leq 2 \\ P+1 & otherwise \end{cases} \quad (6)$$

The mapping from $LBP_{P,R}$ to$LBP_{P,R}^{riu^2}$(superscript "$riu2$"means rotation invariant "uniform" patterns with $U{\leq}2$), which has $P+2$ distinct output values, can be implemented with a lookup table.

4. CLASSIFICATION

The extracted texture features is fed to fusion classifier of nearest neighbor andProbabilistic Neural Network. Introductions on each of these classifiers are given in the flowing subsections.

4.1 Nearest Neighbor (KNN)

One of the least difficult classifiers which we utilized is the K-Nearest Neighbor classifier[29][30]. Term closest can be interpreted as meaning the littlest Euclidean separations in n-dimensional component space. This takes a test highlight in a vector frame, and finds the Euclidean separation amongst this and the vector representation of every preparation case. The preparation test nearest to the test is named as its Nearest Neighbor. Since the prepared specimen in some sense is the one most like our test, it bodes well to apportion its class name to the test.

This endeavors the "smoothness" suspicion that examples close to each other are liable to have the same class

4.2 Probabilistic Neural Networks (PNN)

Probabilistic neural networks ([29] [30]) are forward feed networks built with three layers. They are gotten from Bayes Decision Networks. They prepare rapidly since the preparation is done in one go of every preparation vector, instead of a few. Probabilistic neural systems evaluate the likelihood thickness capacity for every class in light of the preparation samples. The probabilistic neural system utilizes a comparative likelihood thickness capacity. This is computed for every test vector. Vectors must be standardized before contribution to the system. There is an info unit for every measurement in the vector. The info layer is completely associated with the shrouded layer. The shrouded layer has a hub for every order. Each concealed hub ascertains the spot result of the information vector with a test vector subtracts 1 from it and partitions the outcome by the standard deviation squared. The yield layer has a hub for every example order. The entirety for each concealed hub is sent to the yield layer and the most noteworthy qualities wins. The Probabilistic neural system prepares quickly however execution time is moderate and it requires a lot of space in memory.

4.3 Fusion Approach

In this work we study the performance of each classifier individualy. Then we also investigate the suitability of combining two classifiers together. The proposed fusion is a decision level fusion approach using OR operation.

5. RESULTS

In this experimentation, we intend to study the performances of different texture features in script classification and also effect of classifiers. The experimentation has done on database of 30 classes under varying training samples from 40, 60 and 80 percent of database. The results obtained for individual feature, Combination of single feature, Combination of two feature, Combination of three feature and Combination of four feature are respectively tabulated in Table 1, Table2, Table 3,Table 4 and Table 5. In each table, the results are tabulated for each individual classifier and fusion of two classifiers. It shall be noticed that the fusion of all classifiers achieves relatively higher accuracy in all cases.

Table 1: Accuracy of different classifier and their combinations for individual Texture Features

Features	KNN	PNN	Fusion
Gabor	61.11	**57.78**	71.11
	44.76	42.86	56.67
	32.12	30	39.7
CTM	**80**	22.22	**82.22**
	60	19.05	63.81
	36.36	15.76	41.82
LMFilter	17.78	15.56	32.22
	15.24	16.19	29.05
	10.91	10	19.09
Sfilter	18.89	16.67	36.67
	15.24	12.73	29.05
	11.21	8.1	22.42
Rfilter	16.19	10.95	22.22
	13.33	9.09	26.19
	12.22	7.78	20
LBP	18.89	9.09	31.11
	15.24	8.57	24.76
	13.64	7.78	21.21

Table 2: Accuracy of different classifier and their combinations for Combination Two features

Features	KNN	PNN	Fusion
Gabor+CTM	62.22	60	**68.89**
	48.1	45.24	60.95
	33.94	30.61	43.94
Gabor+LMfilter	42.22	32.22	54.44
	33.81	23.81	47.14
	21.82	16.06	32.42
Gabor+Sfilter	52.22	38.89	62.22
	39.05	26.19	51.43
	25.45	15.15	33.94
Gabor+Rfilter	41.11	34.44	50
	29.05	22.38	44.76
	20.61	15.15	33.64
Gabor+LBP	56.67	57.78	72.22
	45.24	43.81	59.52
	32.12	29.39	40.91
CTM+LmFilter	21.11	17.78	40
	20.95	15.24	35.71

	14.55	12.73	25.15
CTM+Sfilter	27.78	20	47.78
	19.05	15.24	32.86
	13.33	13.33	24.24
CTM+Rfilter	22.22	18.89	35.56
	18.1	13.33	30.48
	15.15	11.21	26.06
CTM+LBP	26.67	22.22	42.22
	22.38	23.33	38.1
	20.3	17.58	30.3
LMfilter+Sfilter	18.89	13.33	33.33
	14.76	12.38	23.33
	7.58	6.36	17.27
LMfilter+Rfilter	13.33	11.11	24.44
	13.81	5.15	22.38
	7.88	9.52	17.58
Lmfilter+LBP	20	15.56	31.11
	14.29	12.86	24.29
	12.12	10.61	22.12
Sfilter+Rfilter	13.33	12.22	22.22
	11.9	9.52	20.95
	9.39	8.18	17.27
Sfilter+LBP	18.89	14.44	33.33
	10.95	10.95	22.38
	9.39	10	19.7
Rfilter+LBP	16.67	12.73	29.52
	12.73	10.48	24.24
	12.22	10	23.33

Table 3: Accuracy of different classifier and their combinations for Combination of Three features

Features	KNN	PNN	Fusion
Gabor+CTM+LMfilter	42.22	31.11	53.33
	34.29	25.24	47.14
	23.03	17.27	35.15
Gabor+CTM+Sfilter	53.33	40	65.56
	39.52	28.1	52.86
	26.36	16.06	34.85
Gabor+CTM+Rfilter	43.33	34.44	52.22
	31.43	22.86	47.62
	21.52	15.76	33.03
Gabor+CTM+LBP	57.78	60	**71.11**
	47.14	46.67	60
	33.33	31.21	43.03
CTM+LMfilter+Sfilter	21.11	22.22	37.78

	18.1	17.62	31.9
	8.48	12.73	19.09
CTM+Sfilter+Rfilter	17.78	16.67	32.22
	16.19	11.43	31.9
	10.3	8.79	23.33
CTM+Rfilter+LMfilter	16.67	13.33	27.78
	14.76	12.86	27.62
	7.88	6.36	23.03
Gabor+Rfilter+LMfilter	25.56	15.56	45.55
	23.81	15.24	35.71
	16.36	10.61	27.27
Gabor+Rfilter+Sfilter	32.22	18.88	47.78
	25.24	14.76	42.38
	16.97	9.09	27.58
Gabor+Sfilter+LMfilter	34.44	18.89	47.78
	27.62	19.05	41.43
	18.18	10.61	30.3
Gabor+Rfilter+LBP	42.22	34.44	52.22
	29.52	21.9	45.24
	21.21	14.85	32.12
Gabor+LMfilter+LBP	41.11	32.22	54.44
	31.43	22.38	46.19
	22.42	15.76	33.33
Gabor+Sfilter+LBP	55.56	36.67	66.67
	37.14	25.71	52.38
	24.55	16.06	36.36
CTM+Rfilter+LBP	22.22	17.78	37.78
	21.43	15.24	33.81
	16.36	13.33	29.09
CTM+LMfilter+LBP	24.44	18.89	36.67
	20.95	17.62	36.19
	13.94	14.24	25.45
CTM+Sfilter+LBP	18.89	15.56	42.22
	12.38	13.81	28.57
	12.42	11.21	26.97
Rfilter+Sfilter+LBP	14.44	12.22	25.56
	12.86	9.52	21.9
	10.91	8.48	21.21
Rfilter+LMfilter+LBP	13.33	11.11	31.11
	11.43	9.52	21.43
	6.67	5.45	16.06
Sfilter+LMfilter+LBP	18.89	13.33	30
	15.24	12.86	28.1
	7.58	6.97	17.27

Table 4: Accuracy of different classifier and their combinations for Combination of Four features

Features	KNN	PNN	Fusion
Gabor+CTM+LMfilter+Rfilter	30	24.44	45.56
	25.71	22.86	37.62
	16.36	18.79	28.48
Gabor+CTM+LMfilter+Sfilter	36.67	25.56	50
	29.52	27.14	44.76
	19.39	19.09	32.42
Gabor+LBP+LMfilter+Sfilter	36.67	27.78	55.56
	28.1	24.76	42.86
	20.3	15.45	31.21
Gabor+LBP+LMfilter+Rfilter	25.56	30	47.78
	23.81	23.81	36.67
	16.97	20.3	30.91
Gabor+LBP+Sfilter+Rfilter	32.22	30	45.56
	25.71	26.67	42.38
	17.88	17.88	27.88
Gabor+LBP+CTM+Rfilter	44.44	33.33	54.44
	30.95	22.38	46.19
	22.42	15.15	34.55
Gabor+LBP+CTM+Sfilter	44.44	37.78	**70**
	30.95	27.14	53.33
	22.42	16.36	38.18
Gabor+LBP+CTM+LMfilter	42.22	32.22	58.89
	32.86	24.29	47.14
	23.33	17.58	33.94
Gabor+CTM+Rfilter+Sfilter	35.56	21.11	50
	27.62	15.71	45.24
	18.79	11.21	31.21
Gabor+LMfilter+Rfilter+Sfilter	23.33	25.56	37.78
	23.81	20.48	34.29
	14.24	16.67	25.76
CTM+LMfilter+Rfilter+Sfilter	23.81	11.11	24.44
	16.19	9.05	23.81
	14.55	6.36	19.7
LBP+LMfilter+Rfilter+Sfilter	16.67	8.89	26.67
	13.03	7.62	20.95
	11.9	5.76	18.48
LBP+CTM+ Rfilter+Sfilter	18.89	17.78	35.56
	12.38	11.9	30.48
	10.91	9.09	24.55
LBP+CTM+ Sfilter+LMfilter	24.44	24.44	42.22
	17.62	17.14	31.9

	8.48	12.42	19.39
LBP+CTM+ Rfilter+LMfilter	20	11.11	27.78
	17.62	10.95	27.14
	16.06	6.36	23.03

Table 5: Accuracy of different classifier and their combinations for Combination of Five features

Features	KNN	PNN	Fusion
Gabor+CTM+LMfilter+Rfilter+Sfilter	24.44	25.56	41.11
	24.29	23.81	37.62
	14.24	16.67	26.36
Gabor+LBP+LMfilter+Rfilter+Sfilter	24.44	25.56	40
	23.33	20.95	34.29
	13.94	16.06	25.76
Gabor+CTM+LBP+Rfilter+Sfilter	32.22	28.89	45.56
	27.14	27.14	43.81
	18.79	20.61	31.21
Gabor+CTM+LBP+LMfilter+Rfilter	28.89	25.56	45.56
	24.76	21.43	35.24
	17.58	20.3	30
Gabor+CTM+LBP+LMfilter+Sfilter	37.78	28.89	**55.56**
	29.52	28.57	46.19
	20.91	18.79	34.24
LMfilter+CTM+LBP+Rfilter+Sfilter	19.52	11.11	27.78
	18.89	9.52	26.67
	16.06	6.36	20.3

6. CONCLUSION

In this work we develop a Trilingual script classification system based on the combination of texture features and classifiers. The suitable texture features such as Color Texture Moments, Local binary pattern, LM filter responses, Schmid filter responses, Maximum filter responses and Gabor responses is explored for the purpose of plant classification.It is observed that the fusion classifier achieves relatively good classification accuracy when compared to any other available classifeir. We have created our own database of documents. We conducted experimentation under varying database size and we studied its effect on classification accuracy. The experimental results have shown that the fusion classifier outperforms any individual classifier.

REFERENCES

[1]   Tang Y. Y., Lee S. H and Suen C. Y., 1996. Automatic document processing: a survey. Pattern recognition, Vol. 29, No. 12, pp. 1931-1952.

[2]   Pal U. and Chaudhuri B.B. (2001), 'Automatic identification of English, Chinese, Arabic, Devnagari and Bangla script line', Proceedings of the International Conference on Document Analysis and Recognition, pp. 790-794.

[3]   Pal U., Sinha S. and Chaudhuri B.B. (2003), 'Word-Wise Script Identification From A Document Containing English, Devnagari And Telugu Text', Proceedings of the Document Analysis and Recognition, pp. 213-220.

[4]   Joshi G., Saurabh G. and Jayanthi S. (2006), 'Script Identification from Indian Documents', Proceedings of the Seventh IAPR workshop on Document Analysis Systems, LNCS 3872, pp. 255-267.

[5]   Busch A., Boles W.W. and Sridharan S. (2005), 'Texture for Script Identification', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.27, No.11, pp.1720-1732.

[6]   Spitz A.L. (1997), 'Determination of script, language content of document images', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.19, No.3, pp. 235-245.

[7]   Lu S. and Tan C.L. (2008), 'Script and Language Identification in Noisy and Degraded Document Images', IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 1, pp. 14-24.

[8]   Pal U. and Chaudhuri B.B. (1997), 'Automatic separation of words in multi-lingual multiscript Indian documents', Proceedings of the International Conference on Document Analysis and Recognition, pp. 576-579.

[9]   Pal U. and Chaudhuri B.B. (1999), 'Script Line Separation from Indian Multi-Script Documents,' Proceedings of the International Conference on Document Analysis and Recognition, pp.406- 409.

[10]   Padma M.C. and Nagabhushan P. (2003), 'Identification and Separation of text words of Kannada, Hindi and English languages through discriminating features', Proceedings of the National Conference of Document Analysis and Recognition, pp. 252-260.

[11]   Dhandra B.V., Mallikarjun H., Ravindra H. and Malemath.V.S. (2007), 'Word Level Script Identification in Bilingual Documents through Discriminating Features', Proceedings of International Conference on Signal processing, Communications and Networking, pp. 630-635.

[12]   Dhanya D., Ramakrishnan A.G. and PeetaBasa P. (2002), 'Script Identification In Printed Bilingual Documents,' Sadhana, Vol. 27, Part-1, pp. 73-82.

[13]   Tan C.L., Leong P.Y. and He S. (1999), 'Language Identification in Multilingual documents', Proceedings of the International Symposium on Intelligent Multimedia and Distance Education.

[14]   Pal U. and Chaudhuri B.B. (1999), 'Script Line Separation from Indian Multi-Script Documents,' Proceedings of the International Conference on Document Analysis and Recognition, pp.406- 409.

[15]   Pal U., Sinha S. and Chaudhuri B.B. (2003), 'Word-Wise Script Identification From A Document Containing English, Devnagari And Telugu Text', Proceedings of the Document Analysis and Recognition, pp. 213-220.

[16]   Pati P.B., Sabari Raju S., Pati N. and Ramakrishnan A.G. (2004), 'Gabor filters for document analysis in Indian Bilingual Documents', Proceedings of the International Conference on Intelligent Sensing and Information Processing, pp.123-126

[17]   Pati P.B. and Ramakrishnan A.G. (2006), 'HVS inspired system for script identification in Indian multi-script documents', Seventh IAPR Workshop on Document Analysis Systems, LNCS, Vol. 3872, pp. 380-389.

[18]   SantanuChaudhury, Gaurav Harit, ShekarMadnani, Shet R.B., (2000), Identification of scripts ofIndian languages by Combining trainable classifiers", Proc. of ICVGIP, India.

[19]   Gopal Datt Joshi, Saurabh Garg, and JayanthiSivaswamy, (2006), Script Identification from IndianDocuments, H. Bunke and A.L. Spitz (Eds.): DAS 2006, LNCS 3872, 255–267.

[20] Dhanya D., Ramakrishnan A.G. and Pati P.B., (2002), Script identification in printed bilingualdocuments, Sadhana, vol. 27, 73-82.
[21] Hiremath P S and S Shivashankar, "Wavelet Based Co-occurrence Histogram Features for TextureClassification with an Application to Script Identification in a Document Image", PatternRecognition Letters 29, 2008, pp 1182-1189.
[22] Srinivas Rao Kunte R. and Sudhakar Samuel R.D., (2002), A Neural Approach in On-line ScriptRecognition for Telugu Language Employing Wavelet Features, National Workshop onComputer Vision, Graphics and Image Processing (WVGIP), 188-191.
[23] PeetaBasaPati, S. Sabari Raju, NishikantaPati and A. G. Ramakrishnan, "Gabor filters forDocument analysis in Indian Bilingual Documents", 0-7803-8243-9/04/ IEEE, IClSlP, pp. 123-126, 2004.
[24] Newsam, S. D., and Kamath, C.: Retrieval using texture features in high resolution multi-spectral satellite imagery. In SPIE Conference on Data Mining and Knowledge Discovery: Theory, Tools, and TechnologyVI(2004).
[25] Leung, T., Malik, J.: Representing and recognizing the visual appearance of materials using three-dimensional textons, International Journal of Computer Vision 43(1):29-44, (2001).
[26] Schmid, C.: Constructing models for content-based image retrieval. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, volume 2, pages 39-45(2001).
[27] Geusebroek, M., Smeulders, W. M., Weijer J.: Fast Anisotropic Gauss Filtering. IEEE Transactions on Image Processing, 12(8):938-943(2003).
[28] Varma, M., Zisserman A.: A statistical approach to texture classification from single images, International Journal of Computer Vision: Special Issue on Texture Analysis and Synthesis, 62(1--2):61—81(2005).
[29] Wasserman P D., Advanced Methods in Neural Computing, New York: Van Nostrand Reinhold (1993), pp. 155-61, and pp. 35-55(1993).
[30] Qasem, S. N., and Shamsuddin, S. M. :Generalization Improvement of Radial Basis Function Network Based on Multi-Objective Particle Swarm Optimization, Journal of Artificial Intelligence(2009).