Recognition of Kannada characters extracted from scene images

Deepak Kumar*
Medical Intelligence and Language Engineering (MILE) Laboratory
Department of Electrical Engineering (EE)
Indian Institute of Science (IISc)
Bangalore, INDIA
deepak@ee.iisc.ernet.in

A G Ramakrishnan
Medical Intelligence and Language Engineering (MILE) Laboratory
Department of Electrical Engineering (EE)
Indian Institute of Science (IISc)
Bangalore, INDIA
ramkiag@ee.iisc.ernet.in

ABSTRACT

In this paper, we describe a method for feature extraction and classification of characters manually isolated from scene or natural images. Characters in a scene image may be affected by low resolution, uneven illumination or occlusion. We propose a novel method to perform binarization on gray scale images by minimizing energy functional. Discrete Cosine Transform and Angular Radial Transform are used to extract the features from characters after normalization for scale and translation. We have evaluated our method on the complete test set of Chars74k dataset for English and Kannada scripts consisting of handwritten and synthesized characters, as well as characters extracted from camera captured images. We utilize only synthesized and handwritten characters from this dataset as training set. Nearest neighbor classification is used in our experiments.

Keywords

Binarization, Energy minimization, Discrete cosine transformation, Angular radial transform, English Fonts, Kannada handwritten symbols, Feature vector extraction

1. INTRODUCTION

Character recognition in a non-degraded document image is a solved problem and we have open source and commercially available recognition engines as OCR [27, 31, 32]. However, there is still enough room for development of novel methods to tackle the new problems encountered with scene images. In general, most of the scanned documents are in gray scale format. However, camera captured images have characters with additional colour information and complexity in the form of degradations. Figure 1 shows a few samples of manually cropped characters or symbols from Chars74k dataset [33]. Several approaches exist in the literature for classification of such characters but they have mainly experimented with different classification strategies. Unlike in scanned document images, characters in natural images contain different types of degradations, such as low resolution, uneven illumination, glare, gloss and artistic format. To counter these degradations and to improve recognition, we propose a binarization and feature extraction method for characters in camera captured images.

Figure 1: Kannada and English image samples from Chars74k dataset [33].

Text detection and localization form part of object recognition in computer vision field and have attained a special status in the camera captured document imaging community [16]. Pan et. al proposed a state-of-the-art system for text localization on ICDAR 2005 Robust reading competitions dataset [20, 28]. Kumar and Ramakrishnan have proposed an OTCYMIST algorithm for text segmentation on ICDAR 2011 Robust reading competition [25].In this paper, we have proposed a recognition algorithm for characters manually cropped from scene images. We split our method into binarization, feature extraction and classification stages. In binarization stage, we use an algorithm proposed by Kumar et. al [26] that counters some of the degradations that impact characters in scene images at word level. Similarly, Mishra et. al [24] have used Markov random fields (MRF) to binarize scene images at word level.

*Corresponding author

DAR '12, December 16 2012, Mumbai, IN, India

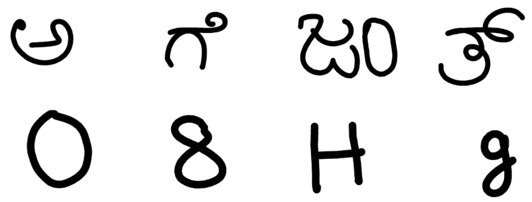

Figure 2: Kannada and English handwritten samples from Chars74k dataset [33].

We use Chars74k dataset [33] for our experiments, which contains both English and Kannada characters. Kannada is one of the ancient Dravidian languages and with a history of more than two thousand years. The training dataset for English has 254 synthetic fonts, and there are four samples in each font, corresponding to normal, bold, italic and bold-italic types. In addition, there are 55 handwritten samples for each of the 62 classes. For Kannada, there are 25 handwritten samples for each of 657 classes. Figure 2 shows a few of the handwritten samples from Chars74k dataset for both Kannada and Roman scripts. Variation in handwritten characters is much higher than in font based characters. Further, training on handwritten characters gives rise to the possibility of recognition, even if handwritten characters are found in camera captured images. The number of classes present in Chars74k *Img* dataset for English and Kannada are 62 and 990, respectively. Thus, the difference in the number of classes is huge, which makes Kannada recognition more complex than that of English.

2. RELATED WORK

The Chars74k dataset was proposed and prepared by de Campos et.al [21, 33]. This dataset includes English and Kannada characters and contains nearly seventy four thousand characters. Most of the English characters are generated artificially using synthetic fonts and the remaining have been obtained by manual cropping from camera captured scene images. Our main interest lies in recognition of Kannada characters from natural images and this is the first public dataset that makes it possible to perform experiments on Kannada character recognition. Font and handwritten samples are used for training and natural images are tested. On the theme of bag-of-visual-words technique, de Campos et. al extract feature vectors to build visual vocabulary. Six different types of local features are calculated namely shape context [8], geometric blur [15], scale invariant feature transform (SIFT) [4], spin image, filter response and patch descriptor. These features generally tend to capture the edge information of a shape to build visual vocabulary. Neumann and Matas [22] use directional feature vectors for training multi-class support vector machines (SVMs) [6] and test only on English *Img* dataset by locating text in a scene image. In other words, they do not directly use the test samples provided in the dataset.

We observe that the feature vectors calculated from the images are local and most of these are dependent on the edges of the training characters. In this paper, we evaluated region based descriptors. An advantage of region based descriptors is that the length of feature vectors used for training is less than that used in other methods.

3. BINARIZATION

The colour images are converted to gray and the gray values of pixels are used for binarization. Kumar et. al [26] have proposed MAPS method to segment manually cropped words from a scene image. Here, we apply MAPS method to segment manually cropped characters. In this approach, the middle row of the image is first binarized using a Min-Max Method.

3.1 Min-Max Method

Locally adaptive methods make use of local statistics to infer the presence or absence of the two classes within the window. If samples from both the classes are present, then the statistics about the two classes are inferred. A priori knowledge of presence of both the classes can also be used for this purpose. In the case of bimodal distributions, the values corresponding to Min and Max belong to different classes, if the spatial window contains samples from both the classes. In the presence of noise, these Min and Max filters get heavily biased. In such cases, statistics about the signal or the two classes cannot be inferred from them. By using a Min-Max filter with a carefully chosen size of window, it is possible to negate the effect of noise. We estimate the maximum values of windows placed to the left and right-side of position i and perform minimum operation to obtain T_{max}. Similarly, maximum operation is performed to obtain T_{min} from the minimum values of left and right-side windows [26].

$$T_{max} = \min(\max_{N_m} L(x_i), \max_{N_m} R(x_i)) \quad (1)$$

$$T_{min} = \max(\min_{N_m} L(x_i), \min_{N_m} R(x_i)) \quad (2)$$

$$L(x_i) = [x_{i-N_m+1}, ..., x_i] \quad (3)$$

$$R(x_i) = [x_i, ..., x_{i+N_m-1}] \quad (4)$$

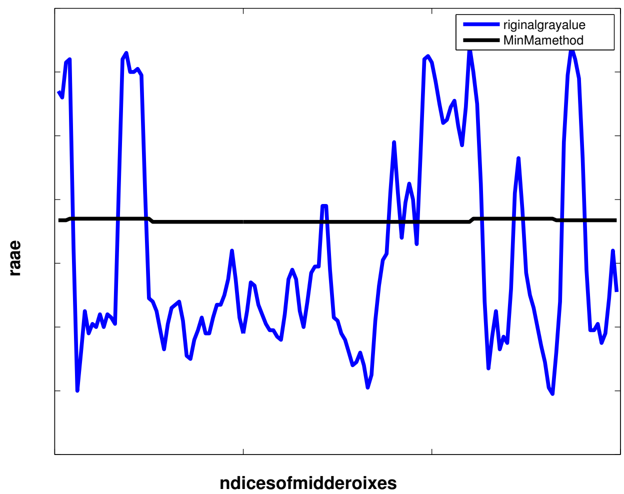

Figure 3: Segmentation of middle row of a degraded document image using Min-Max method

Figure 4: Comparison of outputs of different binarization techniques. (a) A sample image. (b) Classification of pixels using energy minimization function. (c) Otsu's global thresholding. (d) Niblack's local thresholding. (e) Sauvola's and Pietäikinen's local thresholding.

$$T = (T_{min} + T_{max})/2 \tag{5}$$

where, 'N_m' is the size of the moving window obtained as $N_m = min(height, width)$ of the input image. We wrap gray values to compute the threshold at the ends of middle row. Figure 3 shows the plot of middle row gray values and the result of Min-Max method for a sample.

The labels obtained from middle row segmentation are used to estimate the means and variances of the two classes. The mean and variance are calculated as shown:

$$\mu_f = \frac{1}{N_f} \sum_{i \in C_f} x_i \tag{6}$$

$$\sigma_f^2 = \frac{1}{N_f} \sum_{i \in C_f} (x_i - \mu_f)^T * (x_i - \mu_f) \tag{7}$$

where μ_f is the mean of class C_f (foreground), which has N_f labels. σ_f^2 is the variance of class C_f. Similarly, we calculate the mean μ_b and variance σ_b^2 for class C_b (background).

3.2 Min-Cut/Max-Flow

Only the middle row has been considered for segmentation and parameter estimation. Other pixels need to be labeled through classification. We use min-cut/max-flow algorithm for classification of non-middle row pixels. It uses graph cut, which was proposed for image segmentation. Pixels are represented as graph nodes with edges connected to other pixels. Boykov and Kolmogorov have compared min-cut/max-flow algorithms for energy minimization [11]. The energy function of the Potts model, which needs to be minimized, is given as

$$E(L) = \sum_{i \in \mathcal{I}} D_i(L_i) + \sum_{(i,j) \in \mathcal{N}} V_{i,j}(L_i, L_j) \tag{8}$$

where $L = L_i | i \in \mathcal{I}$ is a labeling of the image $\mathcal{I}$, $D(.)$ is a data penalty function, $V_{i,j}$ is an interaction potential, and $\mathcal{N}$ is a set of all pairs of neighboring pixels.

We minimize the energy by incorporating the means and variances as parameters into data penalty and interaction potential functions. The minimum energy results in a binarized image. Figure 4 shows the superior binarization result of the proposed algorithm compared to those of global thresholding Otsu's [1], local thresholding Niblack's [2] and Sauvola's [5] methods for a sample Kannada character image from Chars74k *Img* dataset. For local thresholding algorithms, we have used a window of size 32x32. We observe noise in other methods due to the presence of texture in the background.

4. FEATURE EXTRACTION

We have tried to explore the effectiveness of well established transformations from JPEG and MPEG standards [29, 30]. Discrete cosine transform (DCT) is used in JPEG image compression. DCT features are used in recognition of machine printed Tamil characters by Aparna and Ramakrishnan [10], Kannada characters by Vijay Kumar and Ramakrishnan [9, 12], Arabic word recognition by AlKhateeb et.al [17] and script identification by Pati and Ramakrishnan [18]. Angular radial transform (ART) has been used by Kasar and Ramakrishnan [19] as a feature vector of matching points for mosaicing images. Lavrenko et. al and Rath et. al have used Discrete Fourier transform (DFT) features for word spotting in historical handwritten images [13, 14].

4.1 Discrete cosine transform

Discrete cosine transform was originally suggested for energy compaction [29]. Only a third of the coefficients have significant values, which can be retained as the feature vector of a character. The two dimensional DCT equation is defined as [9],

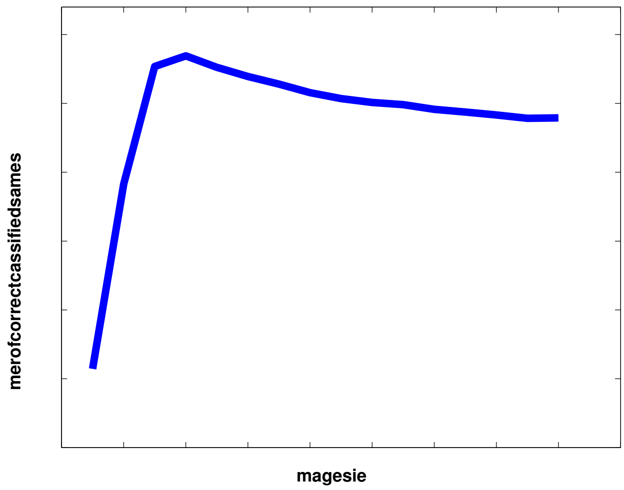

Figure 5: Plot of number of correctly classified samples as a function of normalization size of the image.

Figure 6: Original image of a Kannada character 'th' and images reconstructed using global DCT and block DCT.

$$C(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1}\sum_{y=0}^{M-1} f(x,y)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2y+1)v\pi}{2M}\right] \quad (9)$$

for $0 \leq u \leq (N-1)$,$0 \leq v \leq (M-1)$ and

$$\alpha(k) = \begin{cases} \sqrt{\frac{1}{P}} & k = 0 \\ \sqrt{\frac{2}{P}} & otherwise \end{cases} \quad (10)$$

where, $k = u, v$ and $P = NorM$.

There arises a need to determine the proper size for normalization and further, estimation of feature vectors for training. In our experiment, we explored normalized sizes of binarized image, in multiples of 8, from 8x8 to 128x128. Figure 5 shows the plot of number of correctly classified handwritten Kannada test symbols. A peak occurs at 32x32 and from 40x40 to 128x128, there is minimal variation in the number of correctly classified samples.

After the input image is binarized, we scale the connected components to 32x32 size for unit normalization. We retain 153 DCT coefficients that is equal to 15% significant coefficients in zigzag sequence, similar to JPEG method [29]. We also perform 8x8 block based DCT and retain 15% coefficients for every block. This method of block processing is referred to as 'block DCT' and the other as 'global DCT' in this paper. Figure 6 shows the original image and the two reconstructed images from the retained block and global DCT coefficients.

4.2 Angular radial transform

MPEG-7 standard has evaluated several shape descriptors, which characterize one, two or three dimensional shapes. Any character in a script has some structural information distinct from that of another character in the same script. We obtain the structural information from the characters through shape descriptors [3, 7, 30].

Angular radial transform is a region based descriptor. It is a unitary transform defined on a unit disk and consists of orthonormal sinusoidal basis functions in polar coordinates. ART coefficients are defined as

$$F_{nm} = \int_0^{2\pi}\int_0^1 V_{nm}(\rho,\theta) f(\rho,\theta)\ \rho\ d\rho\ d\theta \quad (11)$$

where F_{nm} is the ART coefficient of order n and m, $f(\rho,\theta)$ is the input image function in polar coordinates and $V_{nm}(\rho,\theta)$ is the ART basis function that is separable along the angular and radial directions.

Accordingly, the basis functions are defined as the product of angular and radial basis functions.

$$V_{nm}(\rho,\theta) = A_m(\theta) * R_n(\rho) \quad (12)$$

Table 1: The cross validation results with different features on Kannada *Hnd* dataset consisting of 657 classes and 25 samples per class.

Feature vector	Recg. rate (%)
Global DCT	**33.3**
Block DCT	33.1
ART	12.7
Shape Context [21]	29.9
Geometric Blur[21]	17.7
SIFT [21]	7.6
Patches [21]	23.0

In order to achieve rotation invariance, an exponential function is used for the angular basis function.

$$A_m(\theta) = (1/2\pi)exp(jm\theta) \quad (13)$$

The radial basis is defined by a cosine function:

$$R_n(\rho) = \begin{cases} 1 & \text{if n = 0} \\ 2\cos(\pi n\rho) & \text{if n} \neq 0 \end{cases} \quad (14)$$

The ART descriptor is the vector of normalized magnitudes of ART coefficients. For scale normalization, the coefficients are divided by the magnitude of ART coefficient of order n= m=0. In the proposed method, individual connected components are rescaled to the fixed size of 45*45 pixels, since the closest natural number for $32 * \sqrt{2}$ is 45. ART coefficients are calculated for order n=5 and m=8, thus giving rise to a feature vector of length 39, excluding F_{00}. These feature vectors are used for training and classification.

5. EXPERIMENTAL RESULTS

We use the training samples of Chars74k dataset to extract feature vectors as described in binarization and feature extraction sections. These reference feature vectors are used to test English *Img* dataset and Kannada *Img* dataset using nearest neighbor classifier. We use the test suite provided by de Campos et. al. Neumann and Matas have used entire image for classification rather than the actual test images in the test suite [22].

5.1 Kannada *Hnd* dataset

Chars74k dataset has 25 samples for each of the 657 classes of Kannada handwritten symbols. These samples are split into 12 samples for training and 13 for testing. Figure 5 shows the plot of correctly classified samples in cross-validation. Nearest neighbor classifier is used for classification. Since the number of training samples available is very small, we cannot resort to training intensive classifiers such as artificial neural network (ANN), hidden Markov models (HMMs) and SVMs. The classification accuracy of our method is tabulated in Table 1 and compared with the reported methods.

5.2 Kannada *Img* dataset

Kannada handwritten symbols are used to extract the feature vectors for training the classifier. The Kannada *Img* test dataset, containing a total of 5135 test samples, is classified using nearest neighbor method. For some strange reason, de Campos et. al have provided only 657 classes in the

Table 2: The classification results on Kannada *Img* 5135 test samples from training Kannada *Hnd* dataset. The number of training classes is 657 and 25 samples per class.

Feature vector	Recg. rate (%)
Global DCT	**11.4**
Block DCT	11.1
ART	1.9
Shape Context [21]	3.5
Geometric Blur [21]	2.8
SIFT [21]	0.3
Patches [21]	0.1

Kannada handwritten dataset and set aside them as training set. On the other hand, they have divided the Kannada *Img* dataset into 990 classes and have called it as the test set. Due to this unusually heavy mismatch in the number of classes between the training and test datasets, the classification results can be expected to be very low. Table 2 shows the classification results on this dataset.

5.3 Results using *Img* dataset for training

English *Img* dataset consists of 12503 characters. de Campos et.al [21] created two training sets, namely Chars74k-5 and Chars74k-15. In Chars74k-5, 5 samples per class are used for training and 15 samples per class in Chars74k-15. Wang et. al generated HOG features from the training samples and used nearest neighbor classifier [23]. Table 3 shows the results on the test set with two separate columns for Chars74k-5 and Chars74k-15. Even for Chars74k-5 dataset, which has very few training samples, DCT based classification accuracy is 2% more than the results reported by Wang et. al. using HOG features. The classification results will be higher for case insensitive recognition, as reported by de Campos et. al [21].

5.4 Results using *Fnt* dataset for training

English *Fnt* dataset consists of 254 different kinds of fonts with the following type faces: normal, bold, italic and italic bold. This synthesized character dataset is used for training. 15 samples per class from English *Img* dataset are used for testing. Table 4 shows the results on this test set. Since individual characters are classified, there is no contextual information. Due to shape topography, some of the numerals and alphabets are classified to other classes. We have reduced the number of classes from 62 to 36 and reported results as class sensitive and insensitive for alphabets.

Table 3: The classification results (%) on English *Img* dataset, using *Img* for training. Nearest neighbor classifier is used since the number of training samples is limited. (5 for Chars74k-5 and 15 for Chars74k-15)

Feature vector	Chars74k-5	Chars74k-15
Global DCT (62-class)	**47.67 ± .65**	57.85
Block DCT (62-class)	**47.71 ± .43**	**58.28**
ART (62-class)	30.82 ± 1.1	41.29
HOG [23]	45.33 ± .99	57.5
MKL [21]	—	55.26
Shape Context [21]	26.1 ± 1.6	34.41
Geometric Blur [21]	36.9 ± 1.0	47.09
Patches [21]	13.7 ± 1.4	21.40
MR8 [21]	6.9 ± 0.7	10.43

Table 4: The classification results on English *Img* dataset, using *Fnt* for training. The number of training classes is 62 and 1016 samples per class. 36-class results are generated by not distinguishing between upper and lower case Roman characters in the output

Feature vector	Recg. rate (%)
Global DCT (36-class)	**74.2**
Block DCT (36-class)	**74.2**
ART (36-class)	61.1
Global DCT (62-class)	66.3
Block DCT (62-class)	66.4
ART (62-class)	49.2
Neumann et.al [22]	71.6
Shape Context [21]	44.8
Geometric Blur [21]	54.3
SIFT [21]	11.1
Patches [21]	7.8

5.5 Results using *Hnd* dataset for training

English *Hnd* dataset consists of 55 samples per class. For this experiment, only this handwritten data is used for training the classifiers. Table 5 shows the classification results for English *Img* dataset. As the number of available training samples per class is less than that in English *Fnt* dataset, the classification result is poor.

6. DISCUSSION

In our method, we have experimented on binarization of scene characters and the size of the feature vector to store and classify the samples. The novelty of our method is in binarization and reduced dimension of the extracted feature vectors. As characters form building blocks of a word in English language, similarly symbols form a word in Kannada language. Apart from cross validation of Kannada handwritten symbols, we also recognize the words in Kannada

Table 5: The classification results on English *Img* dataset, using *Hnd* for training. The number of training classes is 62 and 55 samples per class.

Feature vector	Recg. rate (%)
Global DCT (36-class)	51.6
Block DCT (36-class)	**51.7**
ART (36-class)	34.9
Global DCT (62-class)	46.4
Block DCT (62-class)	44.6
ART (62-class)	26.7
Shape Context [21]	31.1
Geometric Blur [21]	24.6
SIFT [21]	3.1
Patches [21]	1.7

Figure 7: Recognition results using the proposed binarization and classifier. Training set: Chars74k Kannada handwritten training samples. Top: Camera captured Kannada word images. Bottom: Recognized words.

language as shown in Figure 7. In the feature space, the variation of font based classes is less than that of handwritten classes. Due to their spread in the feature vector space, handwritten classes provide advantage in classification of characters or symbols from natural images, which undergo different degradations as well as do not pertain to any specific fonts. In Chars74k dataset, the number of handwritten samples is less; if this is increased, then there will be sufficient data points per class and can avoid curse of dimensionality.

7. CONCLUSION

We have proposed a method to extract feature vectors using DCT and ART as transformation modules after binarizing a natural image. Scene image characters are difficult to binarize than machine printed or handwritten characters. The MAPS method used in binarization stage has improved recognition while testing *Img* dataset as reported in Table 2. The classification accuracy obtained indicates that using DCT as features from binarized image, results in higher recognition rates than other methods. We have made an attempt to build word recognizer from handwritten Kannada symbols. The results (see Fig. 7) are interesting, considering the fact that the number of classes is very huge and that we do not use any lexicon. There is a possibility of pruning the number of classes based on shape topography and consonant-vowel modifiers.

8. REFERENCES

[1] N. Otsu, A Thresholding Selection Method from Gray-level Histogram, *IEEE Trans. SMC*, vol. 9, pp. 62–66, March 1979.

[2] W. Niblack, *An introduction to digital image processing.* New York: Prentice Hall, 1986.

[3] W. Y. Kim and Y. S. Kim, *A new region-based shape descriptor*, TR 15-01, Pisa, December 1999.

[4] D. G. Lowe, Object recognition from local scale-invariant features. Proc. 7th *ICCV*, pp.1150-1157, 1999.

[5] J. J. Sauvola and M. Pietäikinen, Adaptive document image binarization, *Pattern Recognition*, vol. 33, no. 2, pp. 225–236, 2000.

[6] R. O. Duda, P. E. Hart and D. G. Stork., *Pattern classification*, Wiley, 2001.

[7] S. Jeannin, *MPEG-7 Visual part of experimentation Model Version 9.0*, ISO/IEC JTC1/SC29/WG11/N3914, 55th MPEG meeting, Pisa, Italia, January 2001.

[8] S. Belongie, J. Malik and J. Puzicha, Shape matching and object recognition using shape contexts, *IEEE Trans. PAMI*, vol. 24, pp. 509–522, 2002.

[9] B. Vijay Kumar and A. G. Ramakrishnan, Machine Recognition of Printed Kannada Text, Proc. 5th *IAPR Workshop on DAS*, pp. 37–48, Princeton, NJ, August 19-21, 2002.

[10] K. G. Aparna and A. G. Ramakrishnan, A complete Tamil Optical Character Recognition System, Proc. 5th *IAPR Workshop on DAS*, pp. 53–57, Princeton, NJ, August 19-21, 2002.

[11] Y. Boykov and V. Kolmogorov, An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision, *IEEE Trans. PAMI*, vol. 26, no. 9, pp. 1124–1137, September 2004.

[12] B. Vijay Kumar and A. G. Ramakrishnan, Radial basis function and subspace approach for printed Kannada text recognition, Proc. *IEEE ICASSP-04*, vol 5, pp. 321–324, May 17-21, 2004.

[13] V. Lavrenko, T. Rath, and R. Manmatha, Holistic Word Recognition for Handwritten Historical Documents, Proc. *Document Image Analysis for Libraries (DIAL) 2004*, pp. 278–287, 2004.

[14] T. Rath, R. Manmatha, and V. Lavrenko, A Search Engine for Historical Manuscript Images, Proc. *SIGIR '04*, Sheffield, England, pp. 369–376, 2004.

[15] A. C. Berg, T. L. Berg and J. Malik, Shape matching and object recognition using low distortion correspondence, *CVPR*, pp. 26-33, 2005.

[16] S. M. Lucas, Text Locating Competition Results, Proc. 8th *ICDAR 2005*, pp. 80–85, 2005.

[17] J. H. AlKhateeb, J. Ren, J. Jiang and S. S. Ipson, Word-based handwritten Arabic scripts recognition using DCT features and neural network classifier, *IEEE International Multi-Conference on SSD*, 2008.

[18] P. B. Pati and A. G. Ramakrishnan, Word Level Multi-script Identification, *Pattern Recognition Letters*, vol. 29, pp. 1218-1229, 2008.

[19] T. Kasar and A. G. Ramakrishnan, CCD: Connected Component Descriptor for Robust Mosaicing of Camera-Captured Document Images, Proc. 8th *IAPR Workshop on DAS*, Nara, Japan, pp. 480-486, 2008.

[20] Y. F. Pan, X. Hou, C. L. Liu, Text localization in natural scene images based on conditional random field, Proc. 10th *ICDAR 2009*, pp. 6–10, 2009.

[21] T. E. de Campos, B. R. Babu and M. Varma, Character recognition in natural images, *VISAPP*, February 2009.

[22] L. Neumann and J. Matas, A Method for Text Localization and Recognition in Real-World Images, Proc. 10th *ACCV*, pp.770–783, 2010.

[23] K. Wang, S. Belongie, Word Spotting in the Wild, Proc. *ECCV*, pp. 591–604, 2010.

[24] A. Mishra, K. Alahari and C. V. Jawahar, An MRF Model for Binarization of Natural Scene Text, Proc. 11th *ICDAR*, pp. 11–16, 2011.

[25] D. Kumar and A. G. Ramkrishnan, OTCYMIST: Otsu-Canny Minimal Spanning Tree for Born-Digital Images, Proc. 10th *IAPR Workshop on DAS*, pp. 389–393, 2012.

[26] D. Kumar, M. N. Anil Prasad and A. G. Ramakrishnan, MAPS: Midline analysis and propagation of segmentation, Proc. 8th *ICVGIP*, 2012.
[27] "Abbyy Fine reader," http://www.abbyy.com/.
[28] "ICDAR 2005 Competitions," http://http://algoval.essex.ac.uk:8080/icdar2005/index.jsp
[29] "Joint Photographic Experts Group (JPEG)," http://www.jpeg.org/
[30] "Moving Picture Experts Group (MPEG)," http://mpeg.chiariglione.org/
[31] "Nuance Omnipage reader," http://www.nuance.com/.
[32] "Tesseract OCR engine," http://code.google.com/p/tesseract-ocr/.
[33] "The Chars74K image dataset," http://www.ee.surrey.ac.uk/CVSSP/demos/chars74k/