Proceedings of the 2013 International Conference on Machine Learning and Cybernetics, Tianjin, 14-17 July, 2013

GENERATION OF OPTIMAL BINARISATION OUTPUT FROM ANCIENT THAI MANUSCRIPTS ON PALM LEAVES

RAPEEPORN CHAMCHONG[1, 2], CHATKLAW JAREANPON[2], CHUN CHE FUNG[1]

[1]School of Engineering and Information Technology, Murdoch University, Western Australia
[2]School of Information Technology, Faculty of Informatics, Mahasarakham University, Maha Sarakham, Thailand
E-MAIL: rapeeporn.c@gmail.com, chatklaw.j@gmail.com, l.fung@murdoch.edu.au

Abstract:

Recently, several binarisation techniques have been proposed to process different kinds of ancient document images. While many well-known binarisation techniques are particularly suitable for certain types of document images, there is no specific guidelines on the determination of the appropriate type of image degradation, or characteristics of the image. In this paper, a novel method has been proposed to generate the optimal binary image from different binarised outputs from a document image. This approach is based on weight majority vote, and uncertain pixels are then determined based on local areas of the binarised images, by applying iteration of weight majority vote. Experiment over benchmark data set of the Document Image Binarization Contest (DIBCO) 2011 shows that the proposed method provided better performance than most well-known techniques. The proposed method has also been applied to ancient manuscripts on palm leaves from Thailand and this approach provided better results than binarised outputs from original binarisation techniques.

Keywords:

Binarisation; Image segmentation; Document image processing; Historical documents

1. Introduction

Although there exist several binarisation techniques, researchers [1] have proved that there is no single binarisation technique that can be applied effectively to all kinds of digital documents, even in a single application domain. The overall performance of different binarisation systems may vary according to different data sets. Due to this reason, a few researchers pointed to select the optimal binarisation technique or combine binary results from multiple binarisation techniques for an image.

Badekas and Papamakos [2] used a the *Kohonen Self-organizing Map* (KSOM) neural network to learn from multiple results generated by different techniques. This approach worked well with document having complex background, but this method could be time consuming [3]. Gatos et al. [4] proposed the combination of binarisation techniques by using majority vote on the binary input signal from different binarisation results. However, this technique cannot process with an even number of binarisation techniques. The authors compared their methods with six well-known binarisation techniques. F-measure was used to evaluate their proposal and they reported that the combined binary results from multiple binarisation techniques gave the best performance.

Recently, Su et al. [3] proposed a new technique to combine multiple binary results by separating a pixel into three sets, namely, foreground, background, and uncertain pixels. A classifier is then applied to iteratively classify uncertain pixels in to foreground and background by learning from a contrast feature of original grey-scale image. This technique is performed by the combination of two binarisation results. If there are more than two binarisation results, the combined result from first two binarisation results is combined with another binary image as cascading fashion. Evaluation of this technique was compared with the performance due to Otsu's, Sauvola and Pietikainen, Gatos's, Lu's and Su's methods. The four evaluation measures from DIBCO's report were used. This technique is suitable to combine two binarisation results; otherwise, it may be time consuming if more than two results are involved. Due to this reason, this technique will be very time consuming if a large number of binarisation techniques are used.

To provide an optimal binary output for ancient Thai palm leaf manuscripts, the proposed approach is to generate the binary images from multiple binarisation techniques. This approach is based on weight majority vote, and determining uncertain pixels from the information based on local areas of the binary image. This technique could be applied to both even and odd numbers of binary images. The benchmark data set and four evaluation measures from DIBCO 2011's report [5] were used. The proposed method has been applied to generate the optimal binarised output with practical data set which has been collected by Project for Palm Leaf

978-1-4799-0260-6/13/$31.00 ©2013 IEEE

Preservation in Northeastern Thailand Division, Mahasarakham University [6].

In this research, binarisation techniques of the "classical" or the most commonly used approaches based on document binarisation techniques have been used in the proposed method that are *Otsu* (OT) [7] of global threholding technique, *Sauvola and Pietikainen*'s (SAU) technique [8], *Adaptive Logical Level* (ALL) technique [9], *Improvement of Integrated Function* (IIF) algorithm [10], *Background Estimation* (BE) technique [11], and *Local Maximum and Minimum* (LMM) technique [12]. The combination methods of Badekas and Papamakos [2] by using KSOM has also been used to compare with the proposed method, and KSOM method is run by *BDI* application which was distributed by Papamarkos [13].

The remaining of this paper is structured as follows: the binarisation techniques used in the proposed approach are described in the next section followed by the proposed method on the combination of the binarised images explained in Section 3. In following section, experimental results are given, and finally, conclusion is discussed.

2. Combination of Binarised Images

Currently, several binarisation techniques have been proposed and many techniques have claimed good performance on degraded documents. Instead of proposing a new document thresholding technique, this section proposes a combination of multiple existing document thresholding methods for the purpose of producing better binarised results. One way to combine the results is to use majority voting. This technique is explained in next sub-section. In this study, two new techniques for image combination have been proposed that are based on the majority vote concept. These techniques can be applied to both odd and even number of binarised images, and uncertain pixels are considered by using information from the local neighbourhood. These are local adaptation of weight majority vote and local adaption of weighted majority vote that are described in this Sections 2.2 and 2.3, respectively.

2.1. Combined Images Based on Majority Vote

Majority vote employs on labels only, where input label is 1 or 0. The combination result then chooses the largest total vote as shows in Equation (1).

$$\sum_{i=1}^{N} d_{i,J}(x) = \max_{j=0,1} \sum_{i=1}^{N} d_{i,j} \tag{1}$$

However, for a two-class problem, the combined decision will be correct if the number of input class is odd number.

To apply this combination with binarised images, the results from multiple binarisation techniques are represented as $B_i(x,y)$ which is defined as follows.

$$B_i(x,y) = \begin{cases} 0 \ (foreground) \\ 1 \ (background) \end{cases} \tag{2}$$

where $i = 1, 2, \ldots, N$ and $N = 2n+1$.

A combined image $C_B(x,y)$ from N binarisation results is computed from mark as the foreground pixels as follows.

$$C_B(x,y) = \begin{cases} 1 & if \ \sum_{i=1}^{2k+1} B_i(x,y) > n \\ 0 & otherwise \end{cases} \tag{3}$$

However, this technique is not flexible for the number of binarisation results in even number. This study has been improved by considering the number of binarisation results in even number and uncertain pixels. This technique is described in next sub-section.

2.2. Local Adaptation by Weighted Majority Vote

The local adaptation of majority vote by weighted neighbouring pixels is performed as below.

1) Combine binarisation results by using local adaptation of weighted majority vote in each pixel of the image as follows:

1.1) Define normalised weight value of background $\varpi_{1,i}$ and foreground $\varpi_{0,i}$ in neighbouring window of each image.

$$\varpi_{1,i} = \frac{\sum_{y=-m}^{m} \sum_{x=-m}^{m} B_i(x,y)}{M^2} \tag{4}$$

$$\varpi_{0,i} = -(1 - \varpi_{1,i}) \tag{5}$$

where a size of neighbouring window is defined as $M = 2m+1$.

1.2) Adjust weight values in neighbouring window of an input image by using Gaussain distribution mask, $g(x,y)$, as follows:

$$W_i(x,y) = \varpi_i(x,y) \times g(x,y) \tag{6}$$

1.3) Determine uncertain pixel by applying Equation (7).

$$C_U(x,y)=\begin{cases} 0(foreground) & if \dfrac{\sum_{i=1}^{N}\sum_{y=-m}^{m}\sum_{x=-m}^{m} W_i(x,y)}{N}<0 \\ 1\ (background) & if \dfrac{\sum_{i=1}^{N}\sum_{y=-m}^{m}\sum_{x=-m}^{m} W_i(x,y)}{N}>0 \\ -1(uncertain) & otherwise \end{cases} \quad (7)$$

2) Check uncertain pixel from step 1.3) as follows:

2.1) If uncertain pixel occurs, then apply the local adaptation of weighted majority vote in step 1 and the size of neighbouring window is increased by $m=m+1$, and the process is then repeated step 2) until uncertain pixel is set to foreground or background.

2.2) Otherwise, the process is stopped. It is found that the process will repeat 2 or 3 times so the process will be stopped if m is more than 5.

3. Experimental Results

The proposed method has been tested over the benchmark data set from DIBCO 2011's report [5]. A total of 16 images of DIBCO 2011 testing data set consist of eight machine-printed and eight handwritten images resulting in with the ground truth for the evaluation. In addition, the proposed method has also been applied to ancient Thai palm leaf manuscripts.

The evaluation measures consist of F-Measure (*FM*), Peak Signal to Noise Ratio (*PSNR*), Distance Reciprocal Distortion Metric (DRD), and Misclassification Penalty Metric (*MPM*). In particular, these measures are described as follows:

a) FM

$$FM=\frac{2\times RC\times PC}{RC\times PC} \quad (8)$$

where $RC=\dfrac{TP}{TP+FN}$, $PC=\dfrac{TP}{TP+FP}$. *RC* and *PC* refer to the binarisation *Recall* and the binarisation *Precision*, respectively. *TP*, *FP* and *FN* denote the *True Positive*, *False Positive* and *False Negative* values, respectively.

b) PSNR

$$PSNR=10\log_{10}\left(\frac{C^2}{MSE}\right) \quad (9)$$

where Mean Square Error (*MSE*) is calculated from $MSE=\dfrac{\sum_{y=1}^{N}\sum_{x=1}^{M}(b(x,y)-gt(x,y))^2}{M\times N}$, and C is the different value between the foreground and background colours. In this study, all images were converted to binary (0, 1), thus, $C=1$. $b(x,y)$ and $gt(x,y)$ are the pixel of the result image $(M\times N)$ and the ground truth image.

c) DRD

$$DRD=\frac{\sum_{k=1}^{S} DRD_k}{NUBN} \quad (10)$$

where $NUBN$ is the number of the non-uniform (not all black or white pixels) 8×8 blocks in the ground truth image and DRD_k is the distortion of the *k*-th flipped pixel and it is computed using a 5×5 normalized weight matrix (W_{Nm}). DRD_k is the weighted sum of the pixels in the 5x5 block of the ground truth that differ from the centered (k_{th}) flipped pixel at (x,y) in the binarisation result image (*b*) The DRD_k is calculated by following equation.

$$DRD_k=\sum_{i=-2}^{2}\sum_{j=-2}^{2}\left|gt_k(i,j)-b_k(x,y)\right|\times W_{Nm}(i,j) \quad (11)$$

d) MPM

$$MPM=\frac{MP_{FN}+MP_{FP}}{2} \quad (12)$$

where $MP_{FN}=\dfrac{\sum_{i=1}^{N_{FN}} d_{FN}^i}{D}$ and $MP_{FP}=\dfrac{\sum_{j=1}^{N_{FP}} d_{FP}^j}{D}$.

d_{FN}^i and d_{FP}^j denote the distance of the i^{th} false negative and the j^{th} false negative pixel from the contour of the ground truth segmentation. The normalisation factor *D* is the sum over all the pixel-to-contour distances of the ground truth object.

The evaluation results of the DIBCO 2011 data set are shown in Table 1. The output of the proposed method by combining OT, SAU, ALL and IIF technique can provide higher significant performance than the combination method by using KSOM in terms of all metrics.

The proposed method can produce better results than other methods by combining LMM, BE and ALL techniques, which are the best performance method in the DIBCO 2011 contest. The combined image of the proposed method can perform better in terms of F-Measure, PSNR and DRD than the origin binarisation techniques separately. This means a higher precision and better text stroke contour can be obtained after combination.

TABLE 1. EVALUATION RESULT

Binarisation techniques	FM	PSNR	DRD	MPM
OT	77.4611	14.6749	25.8190	24.2516
SAU	73.8697	13.5711	18.1682	22.8970
ALL	80.2972	15.0816	11.9148	16.8095
IIF	80.4826	15.4837	8.0187	6.2410
BE	81.6674	15.5888	11.2353	11.3977
LMM	85.5594	16.7525	5.0244	5.4230
KSOM (OT, SAU, ALL, IIF)	80.1970	14.8095	8.7599	10.1772
Proposed method (OT, SAU, ALL, IIF)	86.1340	16.6267	5.4064	8.0933
Proposed method (BE, LMM)	85.4794	16.7251	5.5841	6.0289
Proposed method (BE, LMM, ALL)	87.2591	17.2200	4.7504	5.9286

The proposed method has been applied to generate the optimal output from ancient Thai manuscripts on palm leaves, and some sample results are shown in Figure 1 and Figure 2. ALL can retain detailed stroke of text and noise may occur surround text, and this technique work well on palm leaf image that has crack or line across text such as in Figure 2(d). BE and LMM can suppress noise more than ALL. However, if image has crack or line across text, these techniques cannot extract text from line, or the line of crack will be happened. Noise surround text can be suppressed by using BE and it is better than LMM. BE may produce more noise in some holes of some characters. From the experiment, LMM can extract text better than BE. In some cases, LMM is unsuitable, if the image has line of crack over text such as in Figure 2(c) and high contrast in some area such as in Figure 1(c). From the sample results of two images, the results show the evidence that the proposed method can provide the optimal output from multiple binarised images.

4. Conclusions

The key contribution of this study is the proposal of a novel method that can be used to combine different binarised outputs form different techniques to produce an optimal output. Instead of designing a new binarisation technique, weight majority vote has been applied to neighbouring window, and determined an uncertain pixel by local adaption with weight majority vote. Experiments over the dataset of recent DIBCO 2011 demonstrated superior performance of the proposed method. Experimental results show that the proposed framework can improve the reported binarisation methods significantly. This method also has been applied to practical documents which are ancient Thai manuscripts from palm leaf. The results illustrated that the proposed method can provide the optimal binarised output.

Acknowledgements

The authors wish to express their thanks and appreciation to members of the Preservation of Palm Leaf Manuscripts Project, Mahasarakham University, Thailand for their support and providing images from their database.

References

[1] E. Badekas and N. Papamarkos, "Document binarisation using Kohonen SOM," IET Image Processing, Vol. 1, pp. 67-84, March 2007.

[2] E. Badekas and N. Papamarkos, "Optimal combination of document binarization techniques using a self-organizing map neural network," Engineering Applications of Artificial Intelligence, Vol. 20, pp. 11-24, February 2007.

[3] B. Su, et al., "Combination of Document Image Binarization Techniques," 2011 International Conference on Document Analysis and Recognition (ICDAR), pp. 22-26, 2011.

[4] B. Gatos, et al., "Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information," 19th International Conference on Pattern Recognition, ICPR 2008., pp. 1-4, 2008.

[5] I. Pratikakis, et al., "ICDAR 2011 Document Image Binarization Contest (DIBCO 2011)," 2011 International Conference on Document Analysis and Recognition (ICDAR), pp. 1506-1510, 2011.

[6] Project for Palm Leaf Preservation in Northeastern Thailand Division. (27 February). Available: http://www.bl.msu.ac.th/2553/english_bl.htm

[7] N. Otsu, "A threshold selection method from gray-level histogram," IEEE Trans. Systems, Man and Cybernetics, Vol. 9, pp. 62-66, 1979.

[8] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," Pattern Recognition, Vol. 33, pp. 225-236, 2000.

[9] Y. Yang and H. Yan, "An adaptive logical method for binarization of degraded document images," Pattern Recognition, Vol. 33, pp. 787-807, 2000.

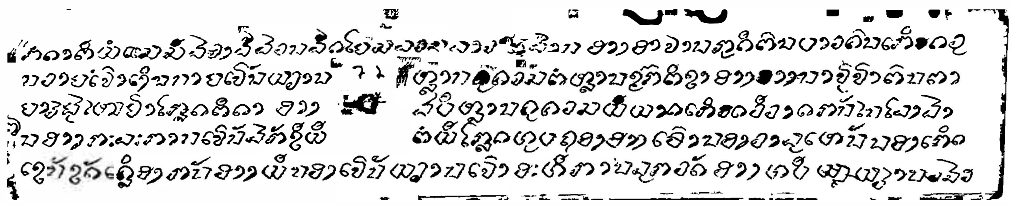
[10] D. Ø. Trier and T. Taxt, "Improvement of “integrated function algorithm” for binarization of document images," Pattern Recognition Letters, Vol. 16, pp. 277-283, 1995.

[11] S. Lu, et al., "Document image binarization using background estimation and stroke edges," Int. J. Doc. Anal. Recognition., Vol. 13, pp. 303-314, 2010.

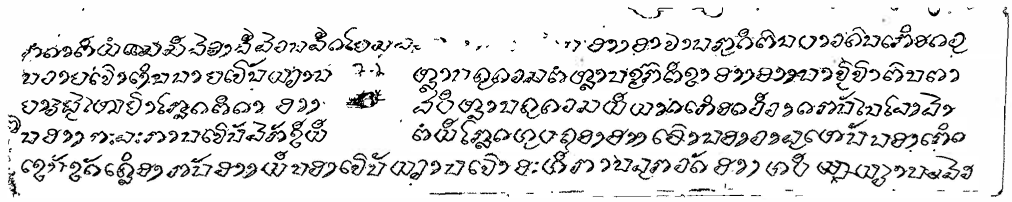
[12] B. Su, et al., "Binarization of historical document images using the local maximum and minimum," Proceedings of the 9th IAPR International Workshop on Document Analysis Systems, Boston, Massachusetts, 2010.

[13] N. Papamarkos. (2007, 13 August). Professor Papamarkos Software. Available: http://www.papamarkos.gr/index/category/134/

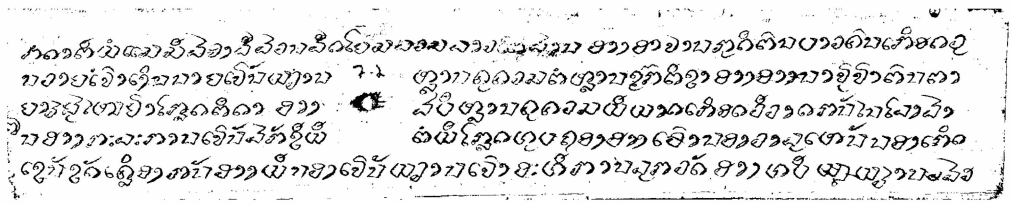
(a) Original



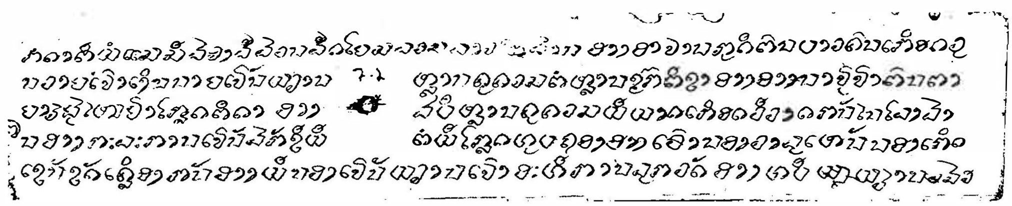
(b) BE



(c) LMM

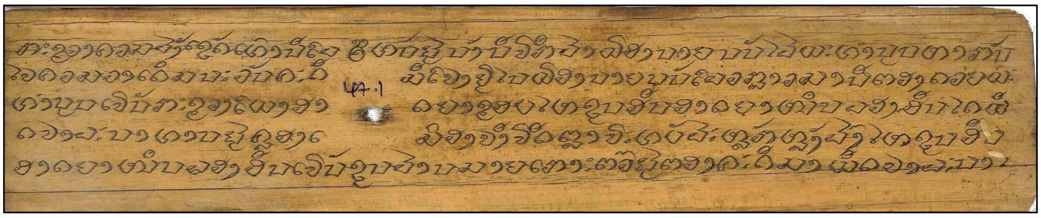


(d) ALL

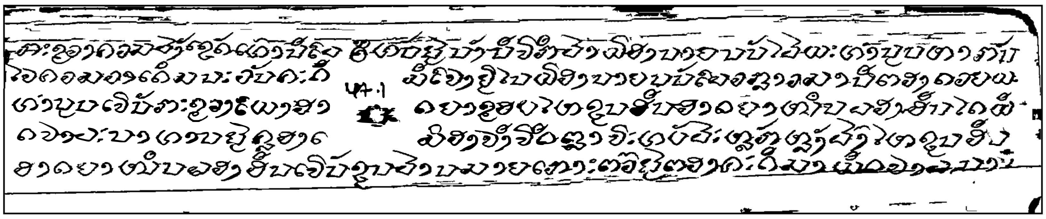


(e) proposed method

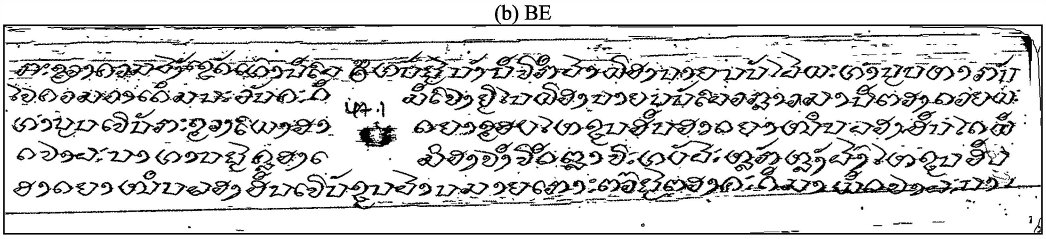
Figure 1. Sample results of ancient Thai palm leaf manuscripts (1)



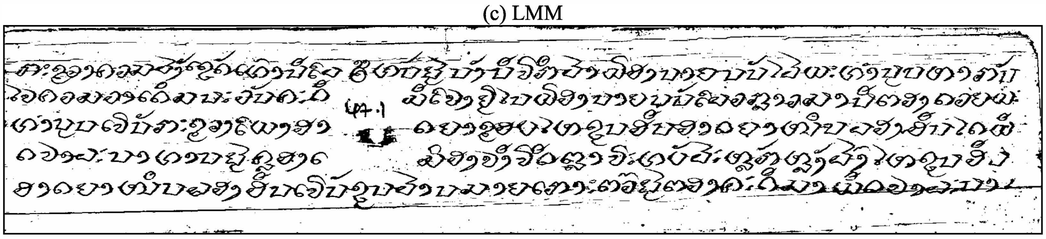
(a) Original



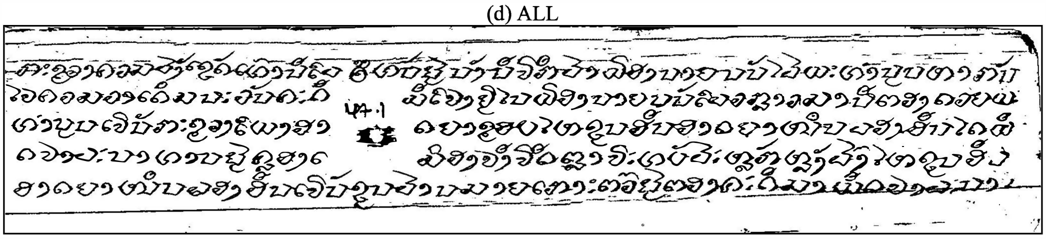
(b) BE



(c) LMM



(d) ALL



(e) proposed method

Figure 2. Sample results of ancient Thai palm leaf manuscript (2)