A Review on Global Binarization Algorithms for Degraded Document Images

Maythapolnun Athimethphat

Department of Business Data Analysis, Faculty of Science and Technology
Assumption University, Bangkok, Thailand
E-mail: <maytha@scitech.au.edu>

Abstract

Several algorithms have previously been proposed for improving the thresholding of degraded document images. No algorithm can solve all types of problems, but some algorithms are better than others for specific situations.

This article reviews global binarization algorithms for improving degraded document images, thus indicating their differences and similarities, and also their advantages and disadvantages. They have been classified into three groups, which are global thresholding, local thresholding and hybrid thresholding. In total, 7 image global threshold binarization algorithms are summarized.

Keywords: *Digital image processing, digitization, thresholding technique.*

1. Introduction

As stated by Khashman and Sekeroglu (2007), "One of the simplest and yet efficient image processing techniques which can be used to separate foreground and background layers of document images is thresholding." Normally, document image analysis uses thresholding as a standard algorithm to change the gray document images to binary form. Document image binarization is very important for old papers to be digitized into digital data.

The degraded document images contain unwanted noises, uneven illumination (shadows), skewed pages, ink seeping, strains, and smear. No standard algorithm performs the best in all degraded document images.

This article reviews the global binarization algorithms for degraded scan images. Ridler and Calvard (1978) developed an algorithm to optimize the process of changing a gray-level image to a bimodal image while retaining the appropriate possible illumination of the image.

Otsu (1979) found a classical algorithm in image binarization which helps reducing a gray-level image to a binary image for classifying foreground and background with a global threshold. He proposed an algorithm that can be applied iteratively to a grayscale histogram of an image for generating threshold candidates.

Pun (1980, 1981) proposed an optimal criterion for image thresholding. This criterion was corrected and improved by Kapur *et al.* (1985). They revised and improved Pun's algorithm by assuming two probability distributions for objects and background as well as maximizing the entropy of the image to obtain the optimal threshold.

Kittler and Illingworth (1986) proposed a minimum error thresholding algorithm that minimizes the probability of classification error by fitting error expression. It is assumed that the image can be characterized by a mixture of two Gaussians distributions of object and background pixels.

Fan *et al.* (1996) proposed a fast entropic technique to obtain a global threshold automatically by reducing complexity in computation.

Portes de Albuquerque *et al.* (2004) proposed an entropic thresholding algorithm that was customized from non-extensive Tsallis entropy concept.

Xiao *et al.* (2008) proposed an entropic thresholding algorithm based on the gray-level spatial correlation (GLSC) histogram. They revised and extended Kapur *et al.*'s algorithm (Kapur *et al.* 1985).

2. Introduction to Binarization Algorithms

Binarization algorithms for document images have been developed for decades and are generally performed at global, local, hybrid or multistage levels.

The following global binarization algorithms are discussed in this review article:

- Algorithm of Ridler and Calvard (1978);
- Algorithm of Kapur *et al*. (1985);
- Algorithm of Fan *et al*. (1996);
- Algorithm of Portes de Albuquerque *et al*. (2004);
- Algorithm of Xiao *et al*. (2008);
- Algorithm of Otsu (1979);
- Algorithm of Kittler and Illingworth (1986).

2.1 Global Binarization Algorithms

Global binarization algorithms use a single global threshold value for a whole document image. The global threshold is used to separate image pixels and background image pixels of objects.

A sample image histogram is shown in Fig. 1. A sample image histogram with global threshold is shown in Fig. 2.

First, in all algorithms considered below, the gray-level histogram is normalized by the following relation:

$$Z = \sum_{i=1}^{L} p_i = \sum_{i=1}^{L} \frac{n_i}{N} = 1, \qquad (1)$$

$$0 \le p_i \le 1 \quad and \quad i = 1,2,...,L,$$

where:

n_i - number of pixels at level i;
N - total number of pixels;
L - number of gray-levels, typically 255.

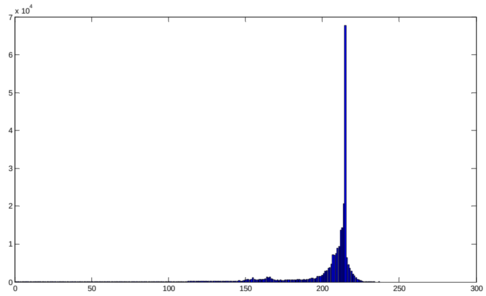

Fig. 1. A sample image histogram.

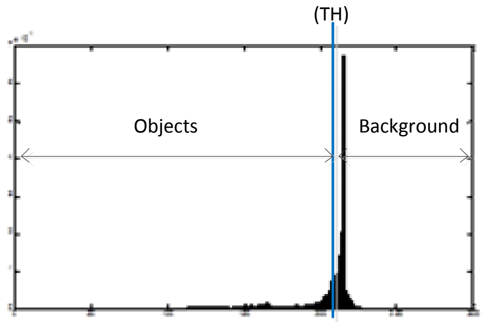

Fig. 2. A sample image histogram with global threshold.

2.1.1 Algorithm of Ridler and Calvard (1978)

In 1978, Ridler and Calvard developed an algorithm to optimize the process of changing a gray-level image to a bimodal image while retaining the appropriate and possible illumination of the image. The purpose of the algorithm of Ridler and Calvard is to evaluate the unique threshold (TH) for any image with bimodal histogram, where TH stands for the threshold value. A sample bimodal image histogram is shown in Fig. 3.

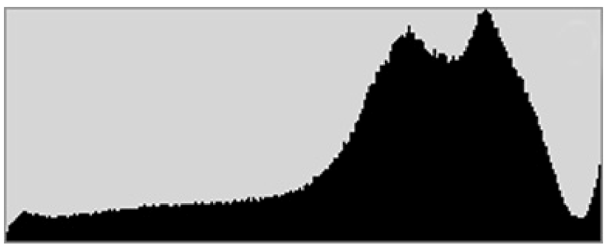

Fig. 3. A sample bimodal image histogram.

Hence, one can divide pixels into background and objects by a threshold at level TH. The probabilities of class occurrence and the class mean level are obtained by:

$$p_{obj} = \sum_{i=0}^{TH} p_i = \omega(TH), \quad (2)$$

$$p_{bg} = \sum_{i=TH+1}^{L} p_i = 1 - \omega(TH), \quad (3)$$

$$p_{obj} + p_{bg} = 1, \quad (4)$$

and

$$\mu_{obj} = \sum_{i=0}^{TH} \frac{ip_i}{p_{obj}} = \frac{\mu(TH)}{\omega(TH)}, \quad (5)$$

$$\mu_{bg} = \sum_{i=TH+1}^{L} \frac{ip_i}{p_{bg}} = \frac{\mu(L) - \mu(TH)}{1 - \omega(TH)}, \quad (6)$$

where $\mu_{bg}(TH)$ and $\mu_{obj}(TH)$ are the mean of background and objects of the histograms separated by threshold (TH), correspondingly. Ridler and Calvard (1978) assumed the optimal threshold as follows:

$$TH_{opt} = \arg\min\left\{\frac{\mu_{bg}(TH) + \mu_{obj}(TH)}{2}\right\}. \quad (7)$$

2.1.2 Algorithm of Kapur *et al*. (1985)

In 1985, Kapur *et al*. revised and improved the algorithm of Pun (1980, 1981) by assuming two probability distributions for objects and background as well as maximizing the entropy of the image to obtain the optimal threshold.

Kapur *et al*. (1985) delineated the between-class entropy of the threshold image as:

$$f_1(TH) = H(0,TH) + H(TH,L). \quad (8)$$

For bi-level thresholding, the optimal threshold is:

$$TH_{optimal} = ArgMax\{f_1(TH)\}. \quad (9)$$

Kapur *et al*. (1985) divides the pixels into two classes by a threshold at level TH. The probabilities of class occurrence and the class mean level are as follows:

$$p_{obj} = \sum_{i=0}^{TH} p_i, \quad (10)$$

$$p_{bg} = \sum_{i=TH+1}^{L} p_i = 1 - p_{obj}, \quad (11)$$

$$p_{obj} + p_{bg} = 1. \quad (12)$$

Thus, the entropies associated with objects and background distributions are:

$$H(0,TH) = -\sum_{i=1}^{TH} \frac{p_i}{p_{obj}} \ln \frac{p_i}{p_{obj}}, \quad (13)$$

$$H(TH,L) = -\sum_{i=TH+1}^{L} \frac{p_i}{p_{bg}} \ln \frac{p_i}{p_{bg}}. \quad (14)$$

For multi-level thresholding, assuming that there are N thresholds $\{TH_1, TH_2, ..., TH_N\}$, the optimal thresholds are chosen by maximizing $f_N(TH_1, TH_2, ..., TH_N)$:

$$\{TH_1^*, TH_2^*, ..., TH_N^*\} = ArgMax\{f_N(TH_1, TH_2, ..., TH_N)\}, \quad (15)$$

$$f_N(TH_1, TH_2, ..., TH_N) = H(0,TH_1) + H(TH_1,TH_2) + ... + H(TH_{N-1},TH_N) = \sum_{i=1}^{N-1} H(TH_i, TH_{i+1}), \quad (16)$$

$$\omega_i = \sum_{j=TH_i}^{TH_{i+1}} p_j, \quad (17)$$

$$H(TH_i, TH_{i+1}) = -\sum_{j=TH_i}^{TH_{i+1}} \frac{p_j}{\omega_j} \ln \frac{p_j}{\omega_j}. \quad (18)$$

The algorithm doesn't involve the image spatial correlation, thus several images with an identical histogram will show the same threshold value. The computational complexity of the algorithm of Kapur *et al*. (1985) is $O(L^2)$.

2.1.3 Algorithm of Fan *et al*. (1996)

In 1996, Fan *et al*. proposed a fast entropic technique to obtain a global threshold automatically by reducing the computational complexity.

The probabilities of class occurrence and the class mean level are obtained by:

$$p_{obj}(TH) = \sum_{i=0}^{TH} p_i, \quad (19)$$

$$p_{bg}(TH) = \sum_{i=TH+1}^{L} p_i = 1 - p_{obj}(TH), \quad (20)$$

$$p_{obj} + p_{bg} = 1, \quad (21)$$

Then for TH+1 one can obtain:

$$p_{obj}(TH+1) = \sum_{i=0}^{TH+1} p_i = p_{obj}(TH) + p_{TH+1}, \quad (22)$$

$$p_{bg}(TH+1) = \sum_{i=TH+2}^{L} p_i = p_{bg}(TH) - p_{TH+1}. \quad (23)$$

Thus, Fan *et al*. (1996) determined two entropies, one for the stationary background, and another for non-stationary classification, which are given by:

$$H_{obj}(TH) = -\sum_{i=1}^{TH} \frac{p_i}{p_{obj}(TH)} \ln \frac{p_i}{p_{obj}(TH)}, \quad (24)$$

$$H_{bg}(TH) = -\sum_{i=TH+1}^{L} \frac{p_i}{p_{bg}(TH)} \ln \frac{p_i}{p_{bg}(TH)}, \quad (25)$$

$$\begin{aligned} H_{obj}(TH+1) &= -\sum_{i=1}^{TH} \frac{p_i}{p_{obj}(TH+1)} \ln \frac{p_i}{p_{obj}(TH+1)} \\ &= -\frac{p_{obj}(TH)}{p_{obj}(TH+1)} \sum_{i=1}^{TH+1} \frac{p_i}{p_{obj}(TH)} \ln\left(\frac{p_i}{p_{obj}(TH)} \frac{p_{obj}(TH)}{p_{obj}(TH+1)} \right) \\ &= \frac{p_{obj}(TH)}{p_{obj}(TH+1)} H_{obj}(TH) - \frac{p_{TH+1}}{p_{obj}(TH+1)} \ln\left(\frac{p_{TH+1}}{p_{obj}(TH+1)} \right) \\ &- \frac{p_{obj}(TH)}{p_{obj}(TH+1)} \ln\left(\frac{p_{obj}(TH)}{p_{obj}(TH+1)} \right), \end{aligned} \quad (26)$$

$$\begin{aligned} H_{bg}(TH+1) &= -\sum_{i=TH+2}^{L} \frac{p_i}{p_{bg}(TH+1)} \ln \frac{p_i}{p_{bg}(TH+1)} \\ &= -\frac{p_{bg}(TH)}{p_{bg}(TH+1)} \sum_{i=TH+2}^{L} \frac{p_i}{p_{bg}(TH)} \ln\left(\frac{p_i}{p_{bg}(TH)} \frac{p_{bg}(TH)}{p_{bg}(TH+1)} \right) \\ &= \frac{p_{bg}(TH)}{p_{bg}(TH+1)} H_{bg}(TH) - \frac{p_{TH+1}}{p_{bg}(TH+1)} \ln\left(\frac{p_{TH+1}}{p_{bg}(TH+1)} \right) \\ &- \frac{p_{bg}(TH)}{p_{bg}(TH+1)} \ln\left(\frac{p_{bg}(TH)}{p_{bg}(TH+1)} \right). \end{aligned} \quad (27)$$

The optimal threshold is:

$$TH_{optimal} = ArgMax\{H_{obj}(TH) + H_{bg}(TH)\}. \quad (28)$$

The computational complexity of the algorithm of Fan *et al*. (1996) is reduced to $O(L)$ by adding the incremental part at each iteration step. This approach is suitable for real-time image segmentation.

2.1.4 Algorithm of Portes de Albuquerque *et al*. (2004)

In 2004, Portes de Albuquerque *et al*. proposed an entropic thresholding algorithm that is customized from non-extensive Tsallis entropy concept.

The probabilities of class occurrence and the class mean level are obtained by:

$$p_{obj} = \sum_{i=0}^{TH} p_i, \quad (29)$$

$$p_{bg} = \sum_{i=TH+1}^{L} p_i = 1 - p_{obj}, \quad (30)$$

$$p_{obj} + p_{bg} = 1. \quad (31)$$

Thus, the entropies associated with objects and background distributions are:

$$H_{obj}(TH) = \frac{1 - \sum_{i=1}^{TH} \left(\frac{p_i}{p_{obj}} \right)^q}{q-1}, \quad (32)$$

$$H_{bg}(TH) = \frac{1 - \sum_{i=TH+1}^{L} \left(\frac{p_i}{p_{bg}} \right)^q}{q-1}. \quad (33)$$

The optimal threshold is:

$$TH_{optimal} = ArgMax\{H_{obj}(TH) + H_{bg}(TH) + (1-q)(H_{obj}(TH)H_{bg}(TH))\}. \quad (34)$$

The computational complexity of the algorithm of Portes de Albuquerque *et al.* (2004) is $O(L^2)$.

2.1.5 Algorithm of Xiao *et al*. (2008)

In 2008, Xiao *et al*. proposed an entropic thresholding algorithm based on the gray-level spatial correlation (GLSC) histogram. As a result, Xiao *et al*. revised and extended the algorithm of Kapur *et al*. (1985).

The GLSC histogram is computed as follows. Let $g(x,y)$ be the number of pixels into small windows which contain $N \times N$ pixels:

$$g(x,y) = \sum_{i=-\frac{N-1}{2}}^{\frac{N-1}{2}} \sum_{-\frac{N-1}{2}}^{\frac{N-1}{2}} \nu(|f(x+i, y+i) - f(x,y)| \le \zeta). \quad (35)$$

The pixel's gray values, $f(x,y)$ and $g(x,y)$ are used to create the GLSC histogram:

$$h(k,m) = \Pr ob(f(x,y) = k \text{ and } g(x,y) = m), \quad (36)$$

where $k \in G$, and $m \in \{1,2,..., N \times N\}$.

The probability function $p(k,m)$ is:

$$p(k,m) = \hat{h}(k,m). \quad (37)$$

The probability distributions associated with object and background are obtained by:

$$G_{obj} = \left\{ \frac{p(0,1)}{p_{obj}}, ..., \frac{p(0, N \times N)}{p_{obj}}, ... , \frac{p(TH,1)}{p_{obj}}, ..., \frac{p(TH, N \times N)}{p_{obj}} \right\}, \quad (38)$$

$$p_{obj} = \sum_{k=0}^{TH} \sum_{m=1}^{N \times N} p(k,m), \quad (39)$$

and

$$G_{bg} = \left\{ \frac{p(TH+1,1)}{p_{bg}}, ..., \frac{p(TH+1, N \times N)}{p_{bg}}, ... , \frac{p(255,1)}{p_{bg}}, ..., \frac{p(255, N \times N)}{p_{bg}} \right\}, \quad (40)$$

$$p_{bg} = \sum_{k=TH+1}^{255} \sum_{m=1}^{N \times N} p(k,m). \quad (41)$$

In the algorithm of Xiao *et al.* (2008), the entropy calculation of elements in the GLSC histogram are weighted by a nonlinear function associated with m and N obtained by:

$$\omega(m,N) = \frac{1 + e^{-9m/N \times N}}{1 - e^{-9m/N \times N}}. \quad (42)$$

Thus, the entropies associated with objects and background distributions are:

$$H_{obj}(TH,N) = -\sum_{k=0}^{TH} \sum_{m=1}^{N \times N} \frac{p(k,m)}{p_{obj}} \ln\left(\frac{p(k,m)}{p_{obj}}\right) \omega(m,N), \quad (43)$$

$$H_{bg}(TH,N) = -\sum_{k=TH+1}^{255} \sum_{m=1}^{N \times N} \frac{p(k,m)}{p_{bg}} \ln\left(\frac{p(k,m)}{p_{bg}}\right) \omega(m,N). \quad (44)$$

Xiao *et al.* (2008) delineated the between-class entropy of the threshold image as:

$$f_1(TH) = H_{obj}(TH,N) + H_{bg}(TH,N). \quad (45)$$

For bi-level thresholding, the optimal threshold is:

$$TH_{optimal} = ArgMax\{f_1(TH,N)\}. \quad (46)$$

The computational complexity of the algorithm of Xiao *et al.* (2008) is $O(N^2 \times L)$. $N = 3$ and $\zeta = 4$ are optimal.

2.1.6 Algorithm of Otsu (1979)

In 1979, Otsu (1979) presented a thresholding algorithm using the histogram shape analysis. This is the most popular global binarization algorithm. This algorithm is very simple. The thresholding of Otsu (1979) shows a favorable performance if the histogram has bimodal distribution. The global threshold (optimum threshold) is selected automatically by a discriminant criterion.

Hence, one can divide pixels into background and objects by a threshold at level TH. The probabilities of class occurrence and the class mean level are obtained by:

$$p_{obj} = \sum_{i=0}^{TH} p_i = \omega(TH), \qquad (47)$$

$$p_{bg} = \sum_{i=TH+1}^{L} p_i = 1 - \omega(TH), \qquad (48)$$

$$p_{obj} + p_{bg} = 1, \qquad (49)$$

and

$$\mu_{obj} = \sum_{i=0}^{TH} \frac{ip_i}{p_{obj}} = \frac{\mu(TH)}{\omega(TH)}, \qquad (50)$$

$$\mu_{bg} = \sum_{i=TH+1}^{L} \frac{ip_i}{p_{bg}} = \frac{\mu(L) - \mu(TH)}{1 - \omega(TH)}, \qquad (51)$$

where:

$$\mu_{Total} = \mu(L) = p_{obj}\mu_{obj} + p_{bg}\mu_{bg} = \sum_{i=0}^{L} ip_i, \qquad (52)$$

is the total mean level of the input gray-level image.

The class variances are given by:

$$\sigma^2_{obj} = \sum_{i=0}^{TH} \frac{(i - \mu_{obj})^2 p_i}{p_{obj}}, \qquad (53)$$

$$\sigma^2_{bg} = \sum_{i=TH+1}^{L} \frac{(i - \mu_{bg})^2 p_i}{p_{bg}}, \qquad (54)$$

$$\sigma^2_{Total} = \sum_{i=0}^{L} (i - \mu(L))^2 p_i = \sigma^2_W + \sigma^2_B, \qquad (55)$$

$$\sigma^2_W = \omega_0\sigma^2_0 + \omega_1\sigma^2_1, \qquad (56)$$

$$\sigma^2_B = \omega_0\omega_1(\mu_1 - \mu_0)^2, \qquad (57)$$

where:

σ^2_0 - variance of object;

σ^2_1 - variance of background;

σ^2_{Total} - total variance of gray-level image;

σ^2_B - variance between-class;

σ^2_W - variance within-class.

It appraises all possible thresholds and finally finds the optimal threshold value that maximizes variance between-class and minimizes variance within-class. The optimal threshold is selected by:

$$\eta = \max_{1 \le TH \le L} \frac{\sigma^2_B}{\sigma^2_W}. \qquad (58)$$

The algorithm of Otsu (1979) can easily fail on blurred images or variety of background images (complex background) because the segmentation of Otsu (1979) is weak with the local variance problem and multimodal distribution as discussed by Lee and Park (1990).

2.1.7 Algorithm of Kittler and Illingworth (1986)

In 1986, Kittler and Illingworth (1986) suggested an algorithm designed for a discriminant object from background in grayscale images. Kittler and Illingworth (1986) assumed that the object and background class conditional probability density function are normal distributions. The histogram is used to classify the error for a mixture of two Gaussians. The algorithm will locate the optimal threshold that minimizes the probability of classification error. This algorithm operates under the assumption of unequal variance Gaussian density function and solves a minimum error Gaussian density-fitting problem.

Hence, one can divide the pixels into background and objects by a threshold at level TH. The probabilities of class occurrence and the class mean level are obtained by:

$$p_{obj} = \sum_{i=0}^{TH} p_i = \omega(TH), \qquad (59)$$

$$p_{bg} = \sum_{i=TH+1}^{L} p_i = 1 - \omega(TH), \qquad (60)$$

$$p_{obj} + p_{bg} = 1, \qquad (61)$$

and

$$\mu_{obj} = \sum_{i=0}^{TH} \frac{ip_i}{p_{obj}}, \quad (62)$$

$$\mu_{bg} = \sum_{i=TH+1}^{L} \frac{ip_i}{p_{bg}}, \quad (63)$$

and, the object and background variance are given by:

$$\sigma_{obj}^2 = \sum_{i=0}^{TH} \frac{(i-\mu_{obj})^2 p_i}{p_{obj}}, \quad (64)$$

$$\sigma_{bg}^2 = \sum_{i=TH+1}^{L} \frac{(i-\mu_{bg})^2 p_i}{p_{bg}}. \quad (65)$$

Kittler and Illingworth (1986) assumed that the Gaussian model $N(\mu_i, \sigma_i^2)$ is characterized by the mean μ_i and by variance σ_i^2 of the Gaussian probability density function (pdf) of a gray-level $Z_k \le TH$ (k = 1, 2,..N), the k^{th} pixel is assigned to Z_k. The Gaussian pdf is known as the Gaussian function:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}. \quad (66)$$

Applying Bays' theorem for continuous variable from the likelihood, one can obtain:

$$p(x \mid \mu_{obj}, \sigma_{obj}) = \frac{1}{\sqrt{2\pi}\sigma_{obj}} e^{-\frac{(x-\mu_{obj})^2}{2\sigma_{obj}^2}}, \quad (67)$$

$$p(x \mid \mu_{bg}, \sigma_{bg}) = \frac{1}{\sqrt{2\pi}\sigma_{bg}} e^{-\frac{(x-\mu_{bg})^2}{2\sigma_{bg}^2}}. \quad (68)$$

Kittler and Illingworth (1986) expressed the thresholding problem as a minimization of a criterion function $J(TH)$, related to the Bayes decision rule for the minimum classification error. Define p_{mix} as a mixture of these two Gaussian distributions by:

$$p_{mix} = (\alpha) p_{bg} + (1-\alpha) p_{obj}, \quad (69)$$

where α is defined by the portions of background and object in the image.

The optimal threshold corresponds to the minimum overlap area between the two normal distributions. The minimum error thresholding method can be calculated as optimal threshold:

$$TH_{opt} = ArgMin[1 + 2[p_{obj} \ln \sigma_{obj}(T) + p_{bg} \ln \sigma_{bg}(T)] - 2[p_{obj} \ln p_{obj} + p_{bg} \ln p_{bg}], \quad (70)$$

$$J(TH) = \hat{P}_{obj}(\ln \hat{\sigma}_{obj} - \hat{P}_{obj}) + \hat{P}_{bg}(\ln \hat{\sigma}_{bg} - \hat{P}_{bg}). \quad (71)$$

Then, the optimal threshold is obtained by

$$TH_{opt} = ArgMin\{J(TH)\}. \quad (72)$$

If the background and objects are quite isolated in term of grey levels, this algorithm is satisfied. Otherwise, this assumption is not true in many situations because sometimes the distributions are not close to the Gaussian model.

3. Conclusion

This survey has concentrated on the algorithms of image binarization based on a global threshold in order to understand parallelisms and a state that involves complementary components between various methods. No single thresholding technique could be claimed as the best method. The thresholding algorithms are simple, fast and easy for digital image segmentation, which is successfully used in a wide binarization of application of digitization.

Several global thresholding techniques and algorithms depend on bimodal histogram (objects and background), and could have a disadvantage due to the double-sided effect in historical document images that are affected by ink bleed through paper and noisy background.

Future work will concentrate on the local binarization and hybrid binarization.

4. Acknowledgement

Appreciation is dedicated to Asst. Prof. Dr. Vorapoj Patanavijit for his encouragement during the preparation of this review article.

The above work was fully supported by the Faculty of Science and Technology, Assumption University of Thailand.

5. References

Fan, J.P.; Wang, R.; Zhang, L.; Xing, D.; and Gan, F., 1996. Image sequence segmentation based on 2D temporal entropic thresholding. Pattern Recognition. Letters 17(10): 1101-7, September.

Kapur, J.N.; Sahoo, P.K.; and Wong, A.K.C. 1985. A new method for gray-level picture thresholding using the entropy of the histogram. Computer Vision, Graphics, and Image Processing 29(3): 273-85, March.

Khashman, A.; and Sekeroglu, B. 2007. Global binarization of document images using a neural network. 2007 Proc. 3rd Int. IEEE Conf. on Signal-Image Technologies and Internet-Based System, Shanghai, China, 16-18 December 2007, pp. 665-72.

Kittler, J.; and Illingworth, J. 1986. Minimum error thresholding. Pattern Recognition 19(1): 41-7, January-February.

Lee, H.; and Park, R.H. 1990. Comments on 'An optimal multiple threshold scheme for image segmentation'. IEEE Transactions on Systems, Man and Cybernetics 20(3): 741- 2, May-June.

Otsu, N. 1979. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics 9(1): 62-6, January.

Portes de Albuquerque, M.; Esquef, I.A. ; and Gesualdi Mello, A.R. 2004. Image thresholding using Tsallis entropy. Pattern Recognition Letters 25(9): 1059-65, July.

Pun, T. 1980. A new method for grey-level picture thresholding using the entropy of the histogram. Signal Processing 2(3): 223-37, July.

Pun, T. 1981. Entropic thresholding: a new approach. Computer Graphics and Image Processing 16(3): 210-39, July.

Ridler, T.W.; and Calvard, E.S. 1978. Picture thresholding using an iterative selection method. IEEE Transactions on Systems, Man and Cybernetics 8(8): 630-2, August.

Xiao, Y.; Cao, Z.G.; and Zhang, T.X. 2008. Entropic thresholding based on gray-level spatial correlation histogram. Proc. 19th Int. Conf. on Pattern Recognition (ICPR 2008), Tampa, FL, USA, 8-11 December 2008, pp. 1-4.