IMAGE MULTI-THRESHOLDING BASED ON SAMPLE MOMENT FUNCTION

Mehmet Sezgin [a], Bülent Sankur[b]

[a]Tübitak Marmara Research Center, Information Technologies Research Institute, Gebze, Kocaeli, Turkey
[b]Boğaziçi University, Electric-Electronic Engineering Department, Bebek , Istanbul, Turkey
sezgin@btae.mam.gov.tr, sankur@boun.edu.tr

ABSTRACT

A new multi-level thresholding method is proposed using Sample Moment Function (SMF). For the binary case the method can make use of the a priori information on whether the target object is found on the darker or lighter part of the scene. Extensive comparisons using a database of NDT images show that the proposed method outperforms in competition 40 other thresholding methods from the literature [1][2], based upon both objective scores and subjective evaluation.

1. INTRODUCTION

Thresholding figures as a practical and effective image segmentation method, especially in those images where pixels belonging to the object are substantially different from the gray levels of the pixels belonging to the background. In vision-based non-destructive testing (NDT) applications, the objects of interest in images are the defective or faulty parts. These regions often differ from their background in one or more physical attribute, such as density, surface reflectivity, absorptivity, texture, temperature etc. For example, in thermography one seeks the spots that are substantially hotter than the background; in textile image the faulty spots have a differing texture density or orientation. Similarly, in ceramic images the surface reflectivity deteriorates in defective regions; in eddy current images the material failure is indicated by a denser current; in endoscopic images the lumen is darker as compared to the surrounding parenchyma etc.

All these applications have in common the fact that background and foreground are distinguishable [3],[4] [5],[6] in the intensity level of one or more attribute. One way to capture the feature contrast between the object and the background is to measure the mean and variance of the object and background regions, as for each setting of the threshold T a different delineation of the foreground and background is obtained.

Thresholding methods may utilize measurement space information (image histogram shape or entropy), object attributes (regularity, compactness), spatial statistics (label continuity), local information etc. Despite the plethora of methods, the quest for new thresholding algorithms that have uniformly good performance over a large set of image instances continues [2]. Thresholding performance, as for any segmentation problem, becomes the determining factor in task-oriented applications, especially because the success of the subsequent operations depend critically upon the outcome of this first step of binarization [3],[4],[5],[6].

In this paper we propose a novel thresholding method that outperforms a large set of competitors. We assess the performance of the proposed bilevel thresholding algorithm against 40 other thresholding methods. The comparison is based on objective segmentation criteria and it uses a database of NDT images. We call it informed thresholding in that the algorithm knows whether the threshold is to be found in the darker or lighter part of the histogram. This is fair in that in specific applications, the expert knows whether the faulty region is denser, hotter, more reflective etc. or vice versa.

2. PROPOSED THRESHOLDING METHOD

The proposed thresholding method is based on *Sample Moment Function, SMF(T)*, which is a function of the threshold value T. Let an image, with $g \in (0,..,L-1)$ gray levels, have a normalized histogram $p(g)$ and cumulative distribution as $P(T)=\sum_{g=0}^{T} p(g)$. The *SMF(T)*, Sample Moment Function, is defined as:

$$SMF(T)=V_0(T)+V_1(T) \tag{1}$$

where background and foreground means $\mu_0(T), \mu_1(T)$ and their sample variances, $V_0(T), V_1(T)$, are computed as follows:

$$\mu_0(T)=\sum_{g=0}^{T} g\,p(g)\ ,\ \mu_1(T)=\sum_{g=T+1}^{L-1} g\,p(g) \tag{2}$$

$$V_0(T)=\sum_{g=0}^{T}\left[\left(g-\frac{\mu_o(T)}{P(T)}\right)^2\frac{p(g)}{P(T)}\right] \tag{3}$$

$$V_1(T)=\sum_{g=T+1}^{L-1}\left[\left(g-\frac{\mu_1(T)}{1-P(T)}\right)^2\frac{p(g)}{1-P(T)}\right] \tag{4}$$

We have observed that the extremum points of *SMF(T)*, yield good candidates for multilevel thresholding in that the resulting map is meaningful. We will denote the set of extrema positions of the *SMF(T)* by the set $\{TC(k)=[T_1,..T_k..,T_N]\}$ which will be tested for thresholding. The number N of extrema is given by the characteristics of the image. To prevent false alarms, we

0-7803-7750-8/03/$17.00 ©2003 IEEE

consider only the extrema that have support K, that is, for a maximum, one has:

$$TC(i) = \arg\{T \mid SMF(T-K) < SMF(T) \text{ and } SMF(T+K) < SMF(T)\}, \forall\, K = 1,2,3 \quad (5)$$

and similarly for a minimum. A good choice for the support of the extremum was found to be K=3.

Bi-level thresholding: In the particular case of bi-level thresholding, only the positions of the first and last extrema are considered. Based on the a priori information, the leftmost extremum or the rightmost extremum is chosen, depending on whether the target object is expected to lie on the dark side or light side. The flow diagram of the proposed thresholding method is given in Fig. 1.

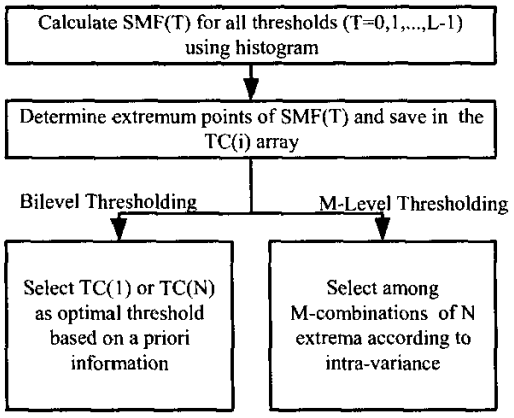


Fig. 1 Flow diagram of the proposed thresholding method.

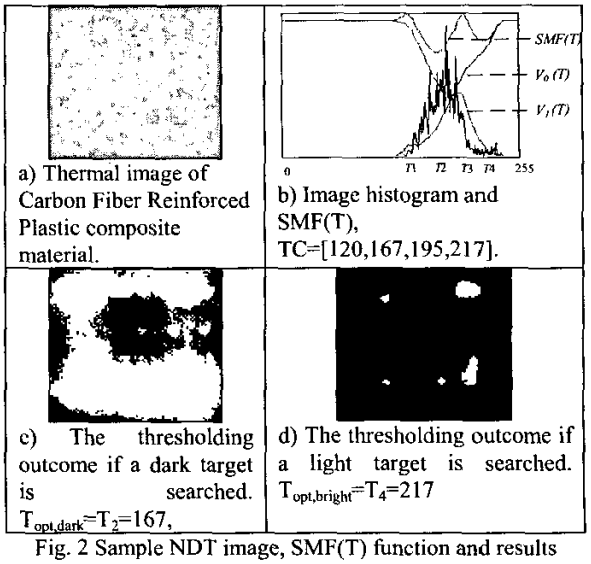
The minimum and maximum points of a sample SMF(T) function is illustrated in Fig. 2. These extremum points are the points where SMF(T) function changes its slope direction (the points TC(i), i = 1..M). To point out the differences between our scheme and the methods that use clustering on the gray level data, we consider Otsu [7], Ramesh [8] and Kittler [9] methods as given in the following equations:

$$T_{Otsu} = \arg\max_T \left\{ \frac{P(T)(1-P(T))[\mu_0(T)-\mu_1(T)]^2}{P(T)\sigma_0^2(T)+(1-P(T))\sigma_1^2(T)} \right\} \quad (6)$$

$$T_{Ramesh} = \arg\min_T [\sum_{g=0}^{T} (\frac{\mu_0(T)}{P(T)} - g)^2 + \sum_{g=T+1}^{L-1} (\frac{\mu_1(T)}{1-P(T)} - g)^2] \quad (7)$$

$$T_{Kittler} = \arg\min_T \left\{ \begin{matrix} P(T)\log\sigma_0(T) + [1-P(T)]\log\sigma_1(T) - \\ P(T)\log P(T) - [1-P(T)]\log(1-P(T)) \end{matrix} \right\} \quad (8)$$

$$T_{proposed} = \arg\, extr_T \left\{ \begin{matrix} \sum_{g=0}^{T} [gP(T) - \mu_0(T)]^2 \frac{p(g)}{P(T)^3}) + \\ \sum_{g=T+1}^{L-1} [g(1-P(T)) - \mu_1(T)]^2 \frac{p(g)}{[1-P(T)]^3} \end{matrix} \right\} \quad (9)$$



a) Thermal image of Carbon Fiber Reinforced Plastic composite material.

b) Image histogram and SMF(T), TC=[120,167,195,217].

c) The thresholding outcome if a dark target is searched. $T_{opt,dark}=T_2=167$,

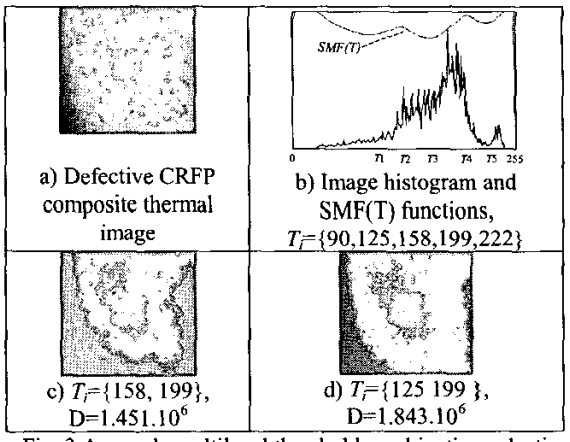
d) The thresholding outcome if a light target is searched. $T_{opt,bright}=T_4=217$

Fig. 2 Sample NDT image, SMF(T) function and results based on dark and bright foreground assumption.

Multi-level thresholding: For M-level multi-thresholding we can still use the extrema of the *SMF(T)* function, for $M \le N+1$, where N is the number of extrema resulting. The best M-1 thresholds can be selected out of $\binom{N}{M-1}$ combinations of the N extrema. Each combination can be ranked according to some performance criterion, like the weighted sum of class variances, D:

$$D = \sum_{k=1}^{N} [P(T_k) - P(T_{k-1})]\sigma_k^2 \quad (10)$$

with $\sigma_k^2 = \sum_{g=T_{k-1}}^{T_k} [g-\mu_k]^2$. A 3-level (M=3) multi-thresholding results are given in Fig. 3 for two different combinations (The lowest two in D).



a) Defective CRFP composite thermal image

b) Image histogram and SMF(T) functions, T_i={90,125,158,199,222}

c) T_i={158, 199}, D=1.451.10^6

d) T_i={125 199 }, D=1.843.10^6

Fig. 3 A sample multilevel threshold combination selection for 3-level thresholding

3. QUANTITATIVE EVALUATION CRITERIA

We considered four complementary segmentation performance metrics, which penalize relatively diverse aspects of thresholding results. These are Misclassification Error (*ME*), Edge-Mismatch (*EMM*), Relative Area Error (*RAE*) and Gray Tone Non-Uniformity (*GTNU*), all normalized to the range [0,1] (best:0, worst:1). The details of these metrics are given in [1],[2]. The performance metric we have used is then the average of these four:

$$AVEM = \frac{1}{4}\{ ME + EMM + GTNU + RAE \} \qquad (11)$$

Quantitative comparison of bi-level thresholding techniques are given for 40 NDT test images drawn from thermal, eddy current, ultrasonic, PCB, cell, textile, tile and material images. The distribution of AVEM performance scores are illustrated in Fig. 4 for the proposed method and its nearest competitor. The ranks of the highest scoring ten methods out of 40 tested are given in Table 1 (ground truths obtained interactively). The proposed method seems to outperform all other methods according to the AVEM criterion. Representative results are illustrated in Table 2.

Table : 1 Ranking and AVEM scores of top thresholding methods averaged over 40 NDT test images.

Rank	Method	Average Score (AVEM)
1	Proposed	0.249
2	Kittler [9]	0.256
3	Kapur [10]	0.261
4	Sahoo [11]	0.269
5	Yen [12]	0.289
6	Lloyd [2]	0.292
7	Otsu [2]	0.318
8	Yanni [2]	0.328
9	Yanowitz [2]	0.339
10	Hertz [2]	0.351
18	Ramesh [8]	0.460

5. CONCLUSION

A multi-thresholding method was introduced based on sample moments SMF. The candidate thresholding values are determined at the extrema points of the SMF. For binarization it suffices to choose either the first or the last extremum, based on a priori information. For more than two levels, a combinatorial search is necessary. Subjective assessments and objective performance scores has revealed that, in a competition with 40 other thresholding methods from the literature [2], run over several NDT images the proposed scheme performs better than all other methods.

6. REFERENCES

[1] M. Sezgin, B. Sankur, Selection of thresholding methods for non destructive testing application, ICIP 2001, Thessaloniki, Greece, 7-10 October.

[2] M. Sezgin, B. Sankur, A survey over image thresholding and quantitative performance evaluation, (accepted) Journal of Electronic Imaging, 2003.

[3] B. S. Wong et al., Thermographic evaluation of defects in composite materials, Insight, 41(8), 1999.

[4] J. Moysan, G. Corneloup, T. Sollier, Adapting an ultrasonic image threshold method to eddy current images and defining a validation domain of the thresholding method, NDT&E Int., 32, 79-84, 1999.

[5] H.J.Krause, et al., Aircraft wheel and fuselage testing with eddy current and SQUID, Insight, 42(3), 2000.

[6] T. Kundu, et al., C-Scan and L-scan generated images of the concrete/GFRP composite interface, NDT&E Int., 32, 61-69, 1999.

[7] N. Otsu, A threshold selection method from gray level histograms, IEEE Tran. Systems, Man, and Cybernetics, SMC-9(1), 62-66, 1979.

[8] N. Ramesh, Yoo, J. H., Sethi, I.K., Thresholding based on histogram approximation, IEE Vision Image and Signal Processing, 142(5), 1995.

[9] J. Kittler, J. Illingworth, Minimum error thresholding, *Pattern Rec.*, 19, 41-47, 1986.

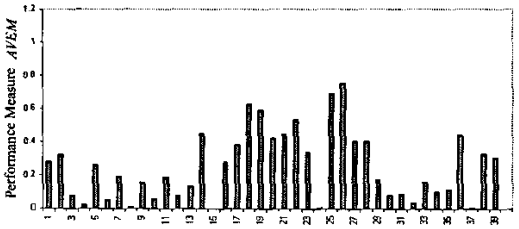
[10] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, GMIP, 29, 273-285, 1985.

[11] P. Sahoo, C. Wilkins, J.Yeager, Threshold selection using Renyi's entropy, Pattern Recognition, 30, 71-84, 1997

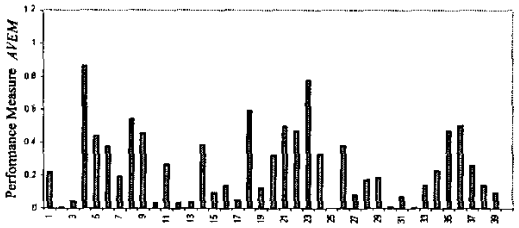
[12] J.C. Yen, F.J. Chang, S. Chang, A new criterion for automatic multilevel thresholding, IEEE Image Processing, IP-4, 370-378, 1995.

[13] V.V. Vinod et. al., A Connectionist approach for gray level image segmentation, 11th International Conference on Pattern Recognition, 489-492, 1992.

[14] M.I Sezan, A peak detection algorithm and its application to histogram-based image data reduction, *CVGIP*, 29, 47-59, 1985.



a) Performance distribution of the proposed method.



b) Performance distribution of Kittler's [9] method

Fig. 4 The *AVEM* scores over the 40 NDT test images

Table 2: Sample results of the proposed bilevel thresholding method .

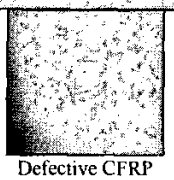
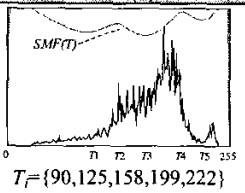
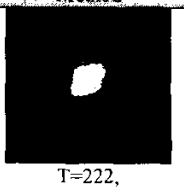
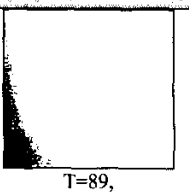
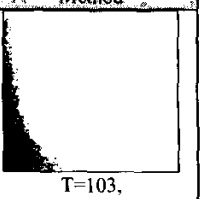
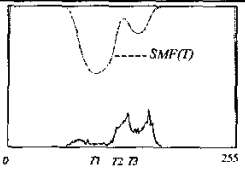
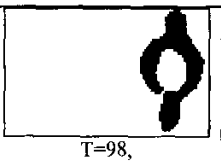
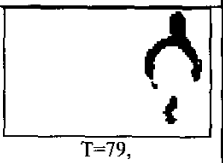
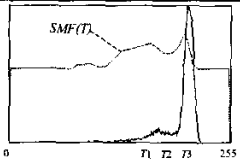
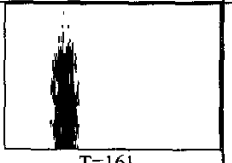
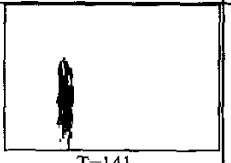
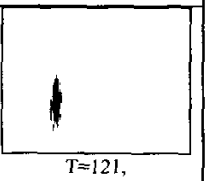
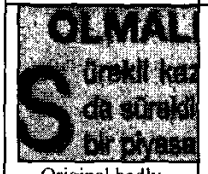
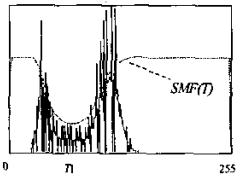
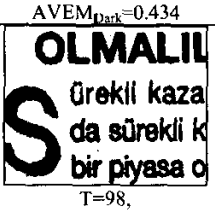
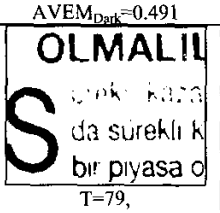
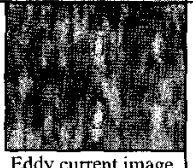
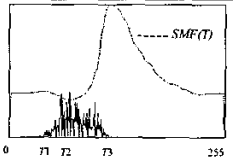
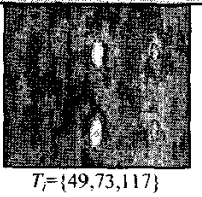
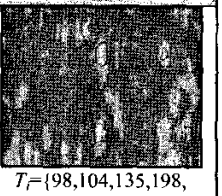
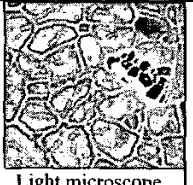
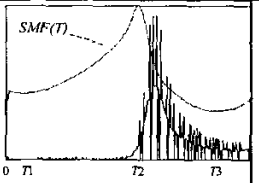
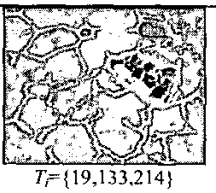
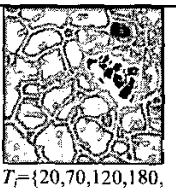
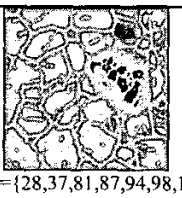
Original Image	Image histogram and SMF(T)	Result of The Proposed Method	Result of Kittler Method	Result of Ramesh Method
 Defective CFRP Composite material thermal band image.	 T_i={90,125,158,199,222}	 T=222, $AVEM_{Bright}$=0.026.	 T=89, $AVEM_{Bright}$ =0.867.	 T=103, $AVEM_{Bright}$ =0.840
 Eddy current image of defective rivet surroundings.	 T_i={98,128,144}	 T=98, $AVEM_{Dark}$ =0.049.	 T=79, $AVEM_{Dark}$ =0.380.	 T=121, $AVEM_{Dark}$ =0.335
 Defective mirror surface	 T_i={161,181,201}	 T=161, $AVEM_{Dark}$=0.434	 T=141, $AVEM_{Dark}$=0.491	 T=121, $AVEM_{Dark}$ =0.521
 Original badly illuminated document image	 T_i={73}	 T=98, $AVEM_{Dark}$ =0.070	 T=79, $AVEM_{Dark}$ =0.233	 T=141, $AVEM_{Dark}$ =0.859

Table 3: Sample results of the proposed multilevel thresholding

Original Image	*SMF(T)*	Result of The Proposed Method	Result of Vinod [13] Method	Result of Sezan [14] Method
 Eddy current image of defective steel part.	 Image histogram (black) and *SMF(T)* function.	 T_i={49,73,117}	 T_i={ 25,75,130,185}	 T_i={98,104,135,198, 209,225}
 Light microscope image of material	 Image histogram (black) and *SMF(T)* function,	 T_i={19,133,214}	 T_i={20,70,120,180, 230}	 T_i={28,37,81,87,94,98,1 08,111,115,178,188 }