International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:1, No:9, 2007

Non-Parametric Histogram-Based Thresholding Methods for Weld Defect Detection in Radiography

N. Nacereddine, L. Hamami, M. Tridi, and N. Oucief

***Abstract*—**In non destructive testing by radiography, a perfect knowledge of the weld defect shape is an essential step to appreciate the quality of the weld and make decision on its acceptability or rejection. Because of the complex nature of the considered images, and in order that the detected defect region represents the most accurately possible the real defect, the choice of thresholding methods must be done judiciously. In this paper, performance criteria are used to conduct a comparative study of four non parametric histogram thresholding methods for automatic extraction of weld defect in radiographic images.

***Keywords*—**Radiographic images, non parametric methods, histogram thresholding, performance criteria.

I. Introduction

THE segmentation constitutes one of the most significant problems in image processing, because the result obtained at the end of this stage strongly governs the final quality of interpretation [1]. The radiographic film images contain weld defects placed in background with different intensities. For such images, intensity is a distinguishing feature that can be used to extract the defects from the background. Therefore, a thresholding technique becomes a strong candidate for efficient radiographic image segmentation.

Thresholding is the process of partitioning pixels in the images into object and background classes based upon the relationship between the gray level value of a pixel and a parameter called the threshold. Because of its efficiency in performance and its simplicity in theory, thresholding techniques have been studied extensively and a large number of thresholding methods have been published [2].

Usually, automatic thresholding approaches are classified into two main groups: global and local. In global methods, a fixed threshold is used for the hole image, whereas in local methods the threshold changes dynamically (local methods are often used when the background is uneven due to the poor illumination condition) and the threshold value is computed for each pixel on the basis of information contained in a local neighborhood of the pixel. In the context of global thresholding many algorithms have been reported in the literature [3].

Manuscript received October 10, 2005. This work was supported in part by the Signal and Image Processing Laboratory of the Welding and NDT Research Centre of Algiers, in collaboration with the Communications and Signal Laboratory of the National Polytechnic School of Algiers.

N. Nacereddine, M. Tridi and N. Oucief are with the Welding and NDT Research Centre, Chéraga, Algiers, 16820 Algeria (corresponding author; e-mail: nacereddine_naf@hotmail.com).

L. Hamami is with the Electronic Department, National Polytechnic School of Algiers.

Finding the correct threshold value to separate an image into desirable foreground and background remains a very important step in image processing process. Furthermore, we are always interested in seeking some special universal algorithm to get the threshold value automatically. Actually, the computer can't "see" the image and figure out the desirable foreground according to the relative background. So the idea is by using a statistics of an image to distinguish the best foreground. Histogram is the most direct and meaningful statistics of an image.

In bi-level thresholding, the histogram of the image is usually assumed to have one valley between two peaks, the peaks representing background and objects, respectively. There are two main approaches to the problem of locating the intensity threshold that ideally represents the bottom of this sometimes elusive histogram valley: parametric and non-parametric techniques. In the non-parametric case, we separate the two gray level classes in an optimum manner according to some posterior criterion, without estimating the parameters of the two distributions. The nonparametric methods are more robust, and usually faster than the parametric. In this study, we implement four non-parametric thresholding methods (Sec. III to Sec. VI) on radiographic film images of welds. The first two methods (Otsu's and Kittler's methods) are based on threshold selection by statistical criteria. Another method (Kapur's method) is based on entropy measurement. The last implemented method (Tsai's method) is based on the preservation of moments between the gray level image and its binarized version.

In Sec. VII we present the comparison methodology and performance criteria. The evaluation results of image thresholding methods, non destructive inspection by radiography are given in Sec. VIII. Finally, Sec. IX draws the main conclusions.

II. Definitions

Let the pixels of the image be represented by L gray levels $\{0,1,2,...,L-1\}$. The number of pixels in level i is denoted by h_i and the total number of pixels is denoted by N. To simplify, the gray level histogram is normalized and regarded as probability distribution function

$$p_i = h_i / N, \quad p_i \geq 0, \sum_{i=0}^{L-1} p_i = 1 \tag{1}$$

Suppose we divide the pixels into two classes C_0 and C_1 by a threshold value at k; C_0 denotes pixels with levels $[0,1,...,k]$ and C_1 denotes pixels with levels $[k+1,...,L-1]$. The probabilities of class occurrences ω, class mean levels μ and class variance for both classes are given by:

$$\omega_0 = \sum_{i=0}^{k} p_i \; ; \; \omega_1 = \sum_{i=k+1}^{L-1} p_i = 1 - \omega_0;$$
$$\mu_0 = \sum_{i=0}^{k} i p_i / \omega_0 \; ; \; \mu_1 = \sum_{i=k+1}^{L-1} i p_i / \omega_1 = \frac{\mu_T - \mu_k}{1 - \omega_0}; \quad (2)$$
$$\sigma_0^2 = \sum_{i=0}^{k} (i - \mu_0)^2 p_i \; ; \; \sigma_1^2 = \sum_{i=k+1}^{L-1} (i - \mu_1)^2 p_i = \sigma_T^2 - \sigma_0^2$$

where

$$\mu_k = \sum_{i=0}^{k} i p_i \; ; \; \mu_T = \sum_{i=0}^{L-1} i p_i \; ; \; \sigma_T^2 = \sum_{i=0}^{L-1} (i - \mu_T)^2 p_i$$

μ_T and σ_T are respectively the total mean and standard deviation.

III. Otsu's Variance Method

Otsu [4] suggested minimizing the weighted sum of within-class variances of the object and background pixels to establish an optimum threshold. Recall that minimization of within-class variances is equivalent to maximization of between-class variance. This method gives satisfactory results for bimodal histogram images.

To measure the thresholding performance, a criterion measure is introduced by Otsu:

$$\eta = \sigma_B^2 / \sigma_T^2 \quad (3)$$

where

$$\sigma_B^2 = \omega_0 (\mu_0 - \mu_T)^2 + \omega_1 (\mu_1 - \mu_T)^2 \quad (4)$$

is the between-class variance which can be simplified to

$$\sigma_B^2 = \omega_0 \omega_1 (\mu_1 - \mu_2)^2 \quad (5)$$

The optimal threshold k_{opt} is given by maximizing η, or equivalently maximizing σ_B^2, since σ_T^2 is independent of k.

$$k_{opt} = \max_k \{\sigma_B^2\} \quad (6)$$

IV. Kittlers and Al's Clustering Algorithm

Kittler and Illingworth [5] describe a method originally designed for segmenting object from background in gray scale images. The procedure views the segmentation as a two class classification problem that can be distinguished based on the gray level histogram of the image. The goal is to find a threshold that minimizes the number of miss-classified pixels. The histogram is viewed as an estimate of the probability density function of a mixture of two clusters. We assume that each component $p_{i,c}$ of the mixture is drawn from a normal distribution with mean μ_i, standard deviation σ_i and *a priori* probability ω_i, so that

$$p_i = \sum_{c=0}^{1} \omega_c p_{i/c} \quad (7)$$

where

$$p_{i/c} = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-(i-\mu_i)^2/2\sigma_i^2} \quad (8)$$

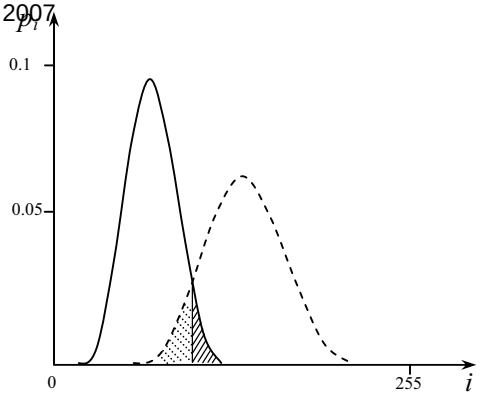

Fig. 1 Classification error for mixture of two Gaussians

The classification error associated with a mixture of two normal distributions is given by the integral of the minimum of the two distributions. The Kittler's and Illingworth's method costs the thresholding problem as a classification problem and seeks the threshold for which the error is minimal.

As shown in Fig. 1, the minimum error threshold can be found by solving the quadratic equation obtained by:

$$\omega_0 \frac{1}{\sqrt{2\pi}\sigma_0} e^{-(i-\mu_0)^2/2\sigma_0^2} = \omega_1 \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(i-\mu_1)^2/2\sigma_1^2} \quad (9)$$

Since the true (μ_i, σ_i, ω_i) are rarely known, they need to be estimated using computationally expensive fitting techniques. Kittller and Illingworth propose a simpler technique for obtaining these estimates. Suppose that the histogram is thresholded at an arbitrary threshold k, then we can model the two resulting populations by a normal density $h_{i/c,k}$ with parameters ($\mu_{i,k}$, $\sigma_{i,k}$, $\omega_{i,k}$) given in the definitions (see Sec. 2.). The probability of level i being correctly classified is given by

$$e_{i,k} = \frac{p_{i/c,k} \omega_{c,k}}{p_i} \quad (10)$$

where $c = \begin{cases} 0 & if \quad i \le k \\ 1 & otherwise \end{cases}$

We wish to find the threshold k that maximizes this function. Since p_i is independent of i and k, it can be safely ignored. Furthermore, since the logarithm is a strictly increasing function, taking the logarithm of both sides will not change the maximizing value. For simplicity, we further multiply by -2 and minimize

$$\varepsilon_{i,k} = \left(\frac{i - \mu_{c,k}}{\sigma_{i,k}} \right)^2 + 2 \ln \sigma_{c,k} - 2 \ln \omega_{c,k} \quad (11)$$

The overall performance is characterized by

$$J(k) = \sum_i p_i \dot{\varepsilon}_{i,k} \quad (12)$$

Writing out *J(k)* gives

$$\sum_{i=0}^{k} p_i \left(\left(\frac{i-\mu_0}{\sigma_0} \right)^2 + 2\ln\sigma_0 - 2\ln\omega_0 \right) + \sum_{i=k+1}^{L-1} p_i \left(\left(\frac{i-\mu_1}{\sigma_1} \right)^2 + 2\ln\sigma_1 - 2\ln\omega_1 \right) \quad (13)$$

which reduces to

$$J(k) = 1 + 2\left(\omega_0 \ln \sigma_0 + \omega_1 \ln \sigma_1\right) - 2\left(\omega_0 \ln \omega_0 + \omega_1 \ln \omega_1\right) \quad (14)$$

Then, the optimal threshold is given by

$$k_{opt} = \min_k \{J(k)\} \quad (15)$$

V. Kapur's Entropy Thresholding

The Kapur's method [6] is based on the entropy theory. It consists in the maximization of the class entropies, which is interpreted as a measure of class compactness and accordingly, of class separability. The suggested probability distributions to represent the foreground and the background respectively are given by:

$$\text{Class } C_0: \quad \frac{p_0}{\omega_k}, \frac{p_1}{\omega_k}, \ldots, \frac{p_k}{\omega_k} \quad (16)$$

$$\text{Class } C_1: \quad \frac{p_{k+1}}{1-\omega_k}, \frac{p_{k+2}}{1-\omega_k}, \ldots, \frac{p_{L-1}}{1-\omega_k} \quad (17)$$

where $\omega_k = \sum_{i=0}^{k} p_i \quad , \quad 1-\omega_k = \sum_{i=k+1}^{L-1} p_i$

The entropies for each class are given by:

$$H(A) = -\sum_{i=0}^{k} \frac{p_i}{\omega_k} \ln \frac{p_i}{\omega_k} \quad (18)$$

$$H(B) = -\sum_{i=k+1}^{L-1} \frac{p_i}{1\text{-}\omega_k} \ln \frac{p_i}{1\text{-}\omega_k} \quad (19)$$

and the total entropy is defined as

$$\Psi(A) = H(A) + H(B) \quad (20)$$

expansion of which leads to:

$$\Psi(k) = \ln\left(\sum_{i=0}^{k} p_i\right) + \ln\left(\sum_{i=k+1}^{L-1} p_k\right) - \frac{\sum_{i=0}^{k} p_i \ln p_i}{\sum_{i=0}^{k} p_i} - \frac{\sum_{i=k+1}^{L-1} p_i \ln p_i}{\sum_{i=k+1}^{L-1} p_i} \quad (21)$$

$$k_{opt} = \max_k \{\Psi(k)\} \quad (22)$$

VI. Tsai's Moment-Preserving Thresholding

Tsai [7] used the preservation of moments to obtain a threshold value without iteration or search. The method also gives representative gray level values for each thresholded class, and the method is easily extended to multi-level thresholding.

Defining m_0 to be *1*, i-th moment m_j of a gray level image *f* may be computed from the normalized histogram h_i

$$m_j = \sum_{i=0}^{L-1} p_i \, i^j \; ; \; j = 0,1,2,3 \quad (23)$$

The image *f* may be seen as a blurred version of an ideal bi-level image *g* with gray levels m_f and m_b ($m_f < m_b$). The method selects a threshold *k* such that if all below-threshold values in *f* are replaced by m_f, and all above-threshold values are replaced by m_b, then the first three moments of *f* are preserved in the unblurred bi-level image *g*. Let ω_0 and ω_1 denote the fractions of the below- threshold and above-threshold pixels in the gray level image. Then the first three moments of the binary image are given by

$$b_j = \omega_0 m_f^j + \omega_1 m_b^j \; ; \; j = 0,1,2,3 \quad (24)$$

Thus, preserving the moments and using the fact that $\omega_0 + \omega_1 = 1 = m_0$, we have a set of four equations:

In the bi-level case, the equations are solved by:

$$c_d = \begin{vmatrix} m_0 & m_1 \\ m_1 & m_2 \end{vmatrix} \quad c_0 = \frac{1}{c_d} \begin{vmatrix} -m_2 & m_1 \\ -m_3 & m_2 \end{vmatrix} \quad c_1 = \frac{1}{c_d} \begin{vmatrix} m_0 & -m_2 \\ m_1 & -m_3 \end{vmatrix}$$

$$m_f = (1/2)\left[-c_1 - (c_1^2 - 4c_0)^{1/2}\right] \quad (25)$$

$$m_b = (1/2)\left[-c_1 + (c_1^2 - 4c_0)^{1/2}\right] \quad (26)$$

$$P_d = \begin{vmatrix} 1 & 1 \\ m_f & m_b \end{vmatrix} \quad (27)$$

$$\omega_0 = (1/P_d) \begin{vmatrix} 1 & 1 \\ m_2 & m_b \end{vmatrix} \quad (28)$$

The optimal threshold *k* is then chosen as the ω_0-tile (or the gray level value closest to the ω_0-tile) of the histogram of *f*.

VII. Thresholding Performance Criteria

In practical thresholding applications, if the thresholding image is complex and the algorithm is fully automatic, the error is inevitable.

The disparity between an actually thresholded image and a correctly/ideally thresholded image (ground-truth of input image) that is the best expected result can be used to assess the performance of algorithms [8]. In the case of the radiographic images of the welded joints, the automated image thresholding encounters difficulties because the object (weld defect) and background gray levels possess substantially overlapping distributions, even resulting in an unimodal distribution. Consequently, misclassified pixels and shape deformations of the object may adversely affect the results of radiographic film interpretation. Therefore, the criteria to assess thresholding algorithms must take into consideration both the noisiness of the segmentation map as well as the shape deformation of weld defects.

To put into evidence the differing performance features of thresholding methods [2], we have used the following three

performance criteria: misclassification error (ME), relative foreground area error (RAE), and region non uniformity (NU). The region non uniformity criterion is not based on ground truth data, but judges the intrinsic quality of the segmented areas. We have adjusted these performance measures so that their scores vary from 0 for a totally correct segmentation to 1 for a totally erroneous case.

A. Misclassification Error

Misclassification error (ME) [9] reflects the percentage of background pixels wrongly assigned to foreground, and conversely, foreground pixels wrongly assigned to background. For the two-class segmentation problem, ME can be simply expressed as:

$$ME = 1 - \frac{|B_O \cap B_k| + |F_O \cap F_k|}{|B_O| + |F_O|} \quad (29)$$

where B_O and F_O denote the background and foreground of the original (ground-truth) image, B_k and F_k denote the background and foreground area pixels in the test image, and $|\,.\,|$ is the cardinality of the set. The ME varies from 0 for a perfectly classified image to 1 for a totally wrongly binarized image.

B. Region Non-Uniformity

This measure [10,8] which does not require ground-truth information, is defined as

$$NU = \frac{|F_k|}{|F_k + B_k|} \frac{\sigma_0^2}{\sigma_T^2} \quad (30)$$

where σ_T^2 represents the variance of the whole image, and σ_0^2 represents the foreground variance. It is expected that a well-segmented image will have a non-uniformity measure close to 0, while the worst case of NU=1 corresponds to an image for which background and foreground are indistinguishable up to second order moments.

C. Relative Foreground Area Error

The comparison of object properties such as area and shape, as obtained from the segmented image with respect to the reference image, has been used in [8] under the name of relative ultimate measurement accuracy (RUMA) to reflect the feature measurement accuracy. We modified this measure for the area feature A as follows:

$$\text{RAE} = \begin{cases} \dfrac{A_O - A_k}{A_O} & \text{if } A_k < A_O \\ \dfrac{A_k - A_O}{A_k} & \text{if } A_k \geq A_O \end{cases} \quad (31)$$

where A_0 is the area of reference image, and A_k is the area of thresholded image. Obviously, for a perfect match of the segmented regions, RAE is zero, while if there is zero overlap of the object area, the penalty is the maximum one.

D. Combination of Measures

To obtain an average performance score from the previous three criteria, we have considered the arithmetic averaging of the normalized scores obtained from the ME, NU, and RAE criteria. In other words, given a thresholding algorithm, for each image the average of ME, NU, and RAE was an indication of its segmentation quality. Thus the performance measure for the i^{th} tested image is written in terms of the scores of the three metrics as:

$$S(i) = \big(ME(i) + NU(i) + RAE(i)\big)/3 \quad (32)$$

The performance criteria measurement for the overall images for a given thresholding method is defined as:

$$S_{all} = \left(\sum_{i=1}^{4} S\big(image(i)\big) \right) / 4 \quad (33)$$

VIII. Experimental Results and Discussion

In order to test the effectiveness of the different 1-D histogram based thresholding methods on real data, a set of four radiographic testing images representing weld defects such as lack of penetration, crack and porosities was used in these experiments. The weld defect images and their corresponding ground truths are shown on Fig. 2.a and Fig.2.b respectively. It is well known in the case of the radiographic images of welded joints that the major part of images presents complicated shape histograms due to several factors [11] such as uneven background illumination. Nevertheless, an appropriate contrast enhancement technique can contribute in the improvement of the thresholding quality. If small contrast between the background and the weld defect regions still remains, the global thresholding will not provide suitable performance, and thus, a thresholding method based on others approaches such as local approach will be recommended. The digitized images of weld defects in Fig. 2.a are thresholded using the optimal threshold value for each of the four methods studied above. The binary images obtained by thresholding are shown in Figs 2.c - 2.f. By using the proposed methods, the optimal thresholds for the images and their corresponding performance measures are reported in Table I.

By examining the thresholding scores we can deduce that in the case of the 2nd, 3rd and 4th images, the best result were provided by the Kapur's method since its performance measure was the least, whereas the better result for the 1st image was provided by Kittler's method. Also, the Kapur's method gives the more accurate results on the overall images (see Table I). Still for overall images, the Kittler's clustering method comes in second position. Not far from Tsai's method, Otsu's algorithm produces comparable results. For the 1st and 3rd images, Kittler's thresholds are not in agreement with those obtained by other methods. This substandard performance can be explained by the fact that these images present large light regions which are represented in the histogram by important thin picks without normal distribution whereas the Kittler's method is well suited for images where classes are normally distributed. So, except for the method of Kapur which outperforms all the other methods, we can get the feeling that none of the other methods is the best or the worst one and they all have fail instance and outstanding result compare to other one in some special cases. From the computational viewpoint, it is worth noting that the four considered methods are fast since they provide the results after only few seconds.

IX. Conclusion

In this study we have investigated experimentally the effectiveness of some 1D histogram based thresholding methods through radiographic images of welds. Four methods, based respectively on the within-class variance minimization, the minimum error clustering, the class entropy maximization and the moment preservation, are implemented on four images representing some types of weld defects. The effectiveness of these methods is measured by the quality of the obtained binary images using the performance criteria which is combination of the misclassification error, the region non-uniformity and the relative foreground area error. After the implementation issues, it is demonstrated that for the considered images generally, the Kapur method proves to be the stronger thresholding tool.

References

[1] L. Soler, G. Malandrin, H. Delinguette, "Segmentation automatique: Application aux angioscanner 3D", *Revue de Traitement de Signal*, vol. 15, 1998.

[2] M. Sezgin, B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation". *Journal of Electronic imaging* 13(1), Jan. 2004, pp. 146–165.

[3] S.U. Lee, S.Y. Chung, and R.H. Park, "A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation", *Computer Vision, Graphics, and Image Processing*, vol. 52, 1990, pp. 171-190.

[4] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms", *IEEE Trans. on Systems, Man, and Cybern.* vol. SMC-9, 1979, pp. 62-66.

[5] J. Kittler and J. Illingworth, "Minimum Error Thresholding", *Pattern Recognition,* 19(1), 1986.

[6] J.N. Kapur, P.K. Sahoo, and A.K.C.Wong, "A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram", *Computer Vision, Graphics, and Image Processing*, vol. 29, 1985, pp. 273-285.

[7] W-H. Tsai, "Moment-Preserving Thresholding: A New Approach", *Computer Vision, Graphics, and Image Processing*, vol. 29, 1985, pp. 377-393.

[8] Y. J. Zhang, "A survey on evaluation methods for image segmentation", *Pattern Recognition*, 29(8), 1996.

[9] W. A. Yasnoff, J. K. Mui, and J. W. Bacus, "Error measures for scene segmentation", *Pattern Recognition,* 9, 1977, pp. 217–231.

[10] M. D. Levine and A. M. Nazif, "Dynamic measurement of computer generated image segmentations", *IEEE Trans. Pattern Anal. Mach. Intell. PAMI-7*, 1985, pp. 155–164.

[11] N. Nacereddine, M. Tridi, S. S. Belaïfa, M. Zelmat, "Weld defect detection in industrial radiography based digital image processing", *International Conference on Signal Processing, ICSP 2004*, Istanbul, Turkey, dec. 2004.

TABLE I
Thresholding Evaluation using Performance Criteria

		Otsu	*Kittler*	*Kapur*	*Tsai*
Ima 1	*Thres*	152	215	173	135
	S	0.1494	**0.0954**	0.1275	0.1560
Ima 2	*Thres*	164	142	110	155
	S	0.4089	0.1677	**0.1385**	0.3068
Ima 3	*Thres*	163	3	79	118
	S	0.4047	0.2483	**0.1912**	0.3507
Ima 4	*Thres*	128	146	102	126
	S	0.3572	0.4553	**0.1832**	0.3476
	S_{all}	0.3300	0.2416	**0.1601**	0,2903

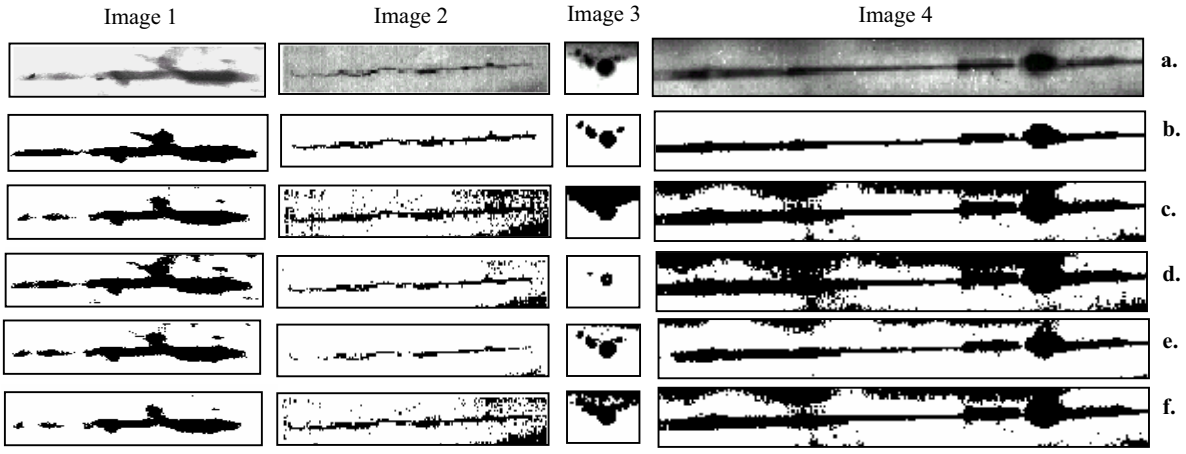

Fig. 2 Thresholding results of weld radiographic images
(a) Weld defect images (b) Ground-truths of a (c) Results of Otsu's thresholding (d) Results of Kittler's thresholding
(e) Results of Kapur's thresholding (f) Results of Tsai's thresholding