International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:7, 2007

Image Thresholding for Weld Defect Extraction in Industrial Radiographic Testing

Nafaâ Nacereddine, Latifa Hamami, and Djemel Ziou

***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 thresholding methods based on gray level histogram, 2-D histogram and locally adaptive approach for weld defect extraction in radiographic images.

***Keywords*—**1D and 2D histogram, locally adaptive approach, performance criteria, radiographic image, thresholding, weld defect.

I. Introduction

CURRENTLY, all the modern industrial realizations cannot be done without the welding which offers undisputable advantages to the welded assemblies such as lightness, high standing to service solicitations, accuracy of execution and esthetic [1]. Nevertheless, this operation is not always perfect because the executed welds can present several types of defects, often caused by manufacturing processes, stresses, environmental changes, a bad choice of the welding procedure, etc. Therefore the inspection operation of welds becomes necessary.

In radiographic testing of welds, the obtained radiographic films are examined by interpreters, of which the task is to detect, recognize and quantify eventual defects and to accept or reject them by referring to the non destructive testing codes and standards. The detection of the defects in a radiogram is sometimes very difficult because of the bad quality of the films, the weld overthickness, the bad contrast, the noise and the weak sizes of defects. The expert often works in extreme cases of the visual system and, that is why the subjectivity in the detection and measurement mechanisms is not negligible.

Manuscript received June 14, 2006.This work was supported 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 and the Department of Mathematics and Computer Science of the University of Sherbrooke.

N. Nacereddine is with the Centre de Recherche en Soudage et Contrôle, Chéraga, Algiers, 16820 Algeria.

(Corresponding author : e-mail: nacereddine_naf@hotmail.com).

L. Hamami is with the Dpt d'Electronique, Ecole Nationale Polytechnique (e-mail : l_hamami@hotmail.com).

Djemel Ziou is with DMI, Faculté des sciences, Université de Sherbrooke, Sherbrooke, Québec, QC J1K 2R1, Canada (e-mail: djemel.ziou@usherbrooke.ca).

Perfect knowledge of the geometry of these weld defects is an important step which is essential to appreciate the quality of the weld [2]. The progresses in computer science and the artificial intelligence techniques have allowed the defect detection and classification to be carried out by using digital image processing and pattern recognition tools, which make the process automatic and more reliable, as it is not a subjective analysis [3].

The segmentation constitutes one of the most significant problems in the image analysis system, because the result obtained at the end of this stage strongly governs the final quality of interpretation [4]. 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 [5] [6].

These methods can be divided, among others, into two categories: global or histogram-based methods and adaptive local methods. Global methods compute a single threshold value for the entire image, and pixels having a gray level value less than the threshold are marked belonging to one class, otherwise the other class. Local methods, on the other hand, compute a threshold value for each pixel on the basis of information contained in a local neighborhood of the pixel.

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. In this paper, statistics based on 1-D and 2-D histograms and local mean and variance of the considered images will be used to distinguish the best foreground representing the most accurately possible, the weld defect region.

Sect. II relates some characteristics about the nature of the radiographic film images and draws the advantages given by the selection of the region of interest (ROI) in radiograms.

In Sect. III, we implement four 1-D histogram-based thresholding methods 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 in this section (Tsai's method) is based on the preservation of moments between the gray level image and its binarized version.

The section IV will be devoted to the cooccurrence matrices and their application in thresholding via local, joint and relative entropies. In Sect. V, locally adaptive methods of Niblack and Sauvola based on gray level statistical properties taken in the neighbourhood of each image pixel are implemented. In Sec. VI we present the comparison methodology and the performance criteria used for the thresholding evaluation. The results obtained from real weld defect images and their discussions are given in Sec. VII. Finally, Sec. VI draws the main conclusions.

II. Digitization and Preprocessing

Generally, the radiographic films are very dark and their density is rather large, therefore an ordinary scanner cannot give a sufficient lighting through a radiogram. Of course, specialized scanners adapted to take high quality copies of radiograms exist, but they are expensive. Here, we have used a scanner AGFA Arcus II, (800 dpi, 256 gray levels). The major part of the radiographic films that we have digitized, were extracted from the standard films provided by International Institute of Welding (IIW). After digitization, the principal characteristics of our images are:

1. Small contrast between the background and the weld defect regions. These last are characterized by unsharpened and blurred edges.
2. Pronounced granularity due to digitization and the type of film used in industrial radiographic testing.
3. Presence of background gradient of image characterizing the thickness variation of the irradiated component part.

For the reasons evoked in the preceding paragraph, it becomes difficult, if not uncertain to detect, during the radiogram visualization, the presence of the small defects and to determine accurately their sizes. That is why, it is often necessary to start with the preprocessing stage in order to reduce or eliminate the noise enclosing in the film and improve its visibility. This procedure permits to obtain an image which would facilitate later the identification of the weld defects being able to be present in the welded joint.

Nevertheless, the first task in image preprocessing is the selection of the region of interest (ROI): the region where they suspect the presence of imperfections. The selection of the ROI saves the operator to make treatments on the useless parts of the image, permitting reduction of the computing time. The second advantage is to save the treatments based on the global approaches to use the irrelevant regions of the image, which can negatively influence the output results. In addition, the limitation of the image to a region of interest (ROI) prevents from the detection of false defects outside the weld. After ROI selection, if necessary, we apply the contrast enhancement of which the goal is to improve the intensity contrast in the input image, highlighting the defect regions whilst leaving the unimportant background regions intact. This enables the defect detection stage to better locate and represent each defect in the image.

III. 1-D Histogram-based Thresholding

A. Definitions

Let the pixels of the image be represented by L gray levels $\{0,1,2,\ldots,L-1\}$. 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 the estimation of 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 and C_1 denote pixels with levels $[0,1,\ldots,k]$ and $[k+1,\ldots,L-1]$, respectively. The probabilities of class occurrences ω, class mean levels μ and class variance for both classes are given by:

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

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.

B. Otsu's Variance Method

Otsu [7] 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. To measure the thresholding performance, a criterion measure is introduced by Otsu:

$$\eta = \sigma_B^2 / \sigma_T^2 \tag{3}$$

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

is the between-class variance which can be simplified to

$$\sigma_B^2 = \omega_0 \omega_1 (\mu_1 - \mu_2)^2 \tag{4}$$

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

$$k_{opt} = \max_k \{\sigma_B^2\} \tag{5}$$

C. Kittler's Clustering Algorithm

In the Kittler's and Illingworth [8] method, the gray level histogram is viewed as an estimate of the probability density function of a mixture of two normal distributions. This method

costs the thresholding problem as a classification problem and seeks the threshold for which the error is minimal. The minimum error threshold can be found by solving the quadratic equation given by:

$$\omega_0 \frac{1}{\sqrt{2\pi}\sigma_0} e^{-(i-\mu_0)^2/2\sigma_0^2} = \omega_1 \frac{1}{\sqrt{2\pi}\sigma_1} e^{-(i-\mu_1)^2/2\sigma_1^2} \quad (6)$$

This amounts to minimizing the criterion:

$$J(k) = 1 + 2(\omega_0 \log \sigma_0 + \omega_1 \log \sigma_1) - 2(\omega_0 \log \omega_0 + \omega_1 \log \omega_1) \quad (7)$$

Then, the optimal threshold is given by

$$k_{opt} = \min_k \{J(k)\} \quad (8)$$

D. Kapur's Entropy Thresholding

The Kapur's method [9] 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:

$$\text{Class } C_0: \quad \frac{p_0}{\omega_0}, \frac{p_1}{\omega_0}, \ldots, \frac{p_k}{\omega_0} \quad (9)$$

$$\text{Class } C_1: \quad \frac{p_{k+1}}{1-\omega_0}, \frac{p_{k+2}}{1-\omega_0}, \ldots, \frac{p_{L-1}}{1-\omega_0} \quad (10)$$

The entropies for each class are given by:

$$H_0 = -\sum_{i=0}^{k} \frac{p_i}{\omega_k} \ln \frac{p_i}{\omega_k} \quad (11)$$

$$H_1 = -\sum_{i=k+1}^{L-1} \frac{p_i}{1\text{-}\omega_k} \ln \frac{p_i}{1\text{-}\omega_k} \quad (12)$$

and the total entropy is defined as

$$\Psi(k) = H_0 + H_1 \quad (13)$$

The optimal is then found by:

$$k_{opt} = \max_k \{\Psi(k)\} \quad (14)$$

E. Tsai's Moment-Preserving Thresholding

Tsai [10] 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*, the i^{th} moment m_j of a gray level image *f* may be computed as

$$m_j = \sum_{i=0}^{L-1} p_i \, i^j \; ; \; j = 0,1,2,3 \quad (15)$$

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

$$b_j = \omega_0 m_f^j + \omega_1 m_b^j \; ; \; j = 0,1,2,3 \quad (16)$$

Thus, preserving the moments and using the fact that $\omega_0 + \omega_1 = 1 = m_0$, we have a set of four equations which, in the bi-level case, are solved by:

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

$$m_f = (1/2)\left[-c_1 - (c_1^2 - 4c_0)^{1/2}\right] \quad (17)$$

$$m_b = (1/2)\left[-c_1 + (c_1^2 - 4c_0)^{1/2}\right] \quad (18)$$

$$\omega_0 = (1/P_d) \begin{vmatrix} 1 & 1 \\ m_2 & m_b \end{vmatrix} \quad (19)$$

where $P_d = \begin{vmatrix} 1 & 1 \\ m_f & m_b \end{vmatrix}$

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

IV. 2-D Histogram-based Tresholding

A. Cooccurrence Matrix and Quadrants

The cooccurrence matrix is a 2-D histogram that describes the occurrence of pairs of pixels that are separated by a certain space vector $\vec{d}$ whose Cartesian coordinates are Δx and Δy and whose polar coordinates are *d* (distance) and θ (orientation). Given a digitized image *I* of size $P \times Q$ with *L* gray levels $G=\{0, 1, 2, \ldots, L-1\}$ used to represent the image *I*. So, for two pixels $(x_1, y_1; i)$ and $(x_1+\Delta x, y_2+\Delta y; j)$ with gray levels *i* and *j*, the cooccurrence matrix $C_{d,\theta}$ is defined as $C_{d,\theta} = [c_{i,j}]_{L \times L}$ with:

$$C_{\theta,d}(i,j) = card\left\{\begin{array}{l}((x_1,y_1),(x_1+\Delta x, y_1+\Delta y)) \in (P\times Q)\times(P\times Q) \\ \Delta x = d\sin\theta, \Delta y = d\cos\theta, \\ I(x_1,y_1) = i, I(x_1+\Delta x, y_1+\Delta y) = j\end{array}\right\} \quad (20)$$

We take after $C_{d,\theta}(i,j)=c_{i,j}$. The probability of cooccur-rence $p_{i,j}$ of gray level *i* and *j* can be written as follows:

$$p_{i,j} = \frac{c_{i,j}}{\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} c_{i,j}}, \; 0 \le p_{i,j} < 1 \quad (21)$$

Let *k* ($k \in G$) be a threshold which partitions the co-occurrence matrix, into four quadrants, *A*, *B*, *C* and *D*, as shown in Figure 1. Pal and al. [11] are separated the four quadrants into two types. One is objects (foreground, those

pixels with gray levels above the threshold) and the other is the background (those pixels with gray levels below); the quadrants *C* and *A* correspond to them, respectively.

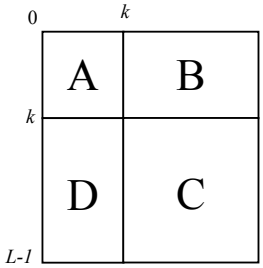

Fig. 1 Quadrants of co-occurrence matrix

However, inversely in our application, the quadrants *A* and *C* correspond to the object and the background respectively, because the majority of weld defects are appeared as dark spots on a radiographic film. Another is the transitions across the boundaries of background and objects; that is, quadrants *B* and *D*. The probabilities associated with each quadrant are defined as follows [12]:

$$P_A(k)=\sum_{i=0}^{k}\sum_{j=0}^{k}p_{i,j}\quad;\quad P_B(k)=\sum_{i=0}^{k}\sum_{j=k+1}^{L-1}p_{i,j} \tag{22}$$

$$P_C(k)=\sum_{i=k+1}^{L-1}\sum_{j=k+1}^{L-1}p_{i,j}\;;\;P_D(k)=\sum_{i=k+1}^{L-1}\sum_{j=0}^{k}p_{i,j} \tag{23}$$

Normalizing the probabilities within each individual quadrant, the cell probabilities are defined as follows:

$$p_{i,j}^{A}=\frac{p_{i,j}}{P_A}=\frac{c_{i,j}\,/\left(\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}c_{i,j}\right)}{\sum_{i=0}^{k}\sum_{j=0}^{k}c_{i,j}\,/\left(\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}c_{i,j}\right)}=\frac{c_{i,j}}{\sum_{i=0}^{t}\sum_{j=0}^{t}c_{i,j}},\quad 0\le i,j\le k \tag{24}$$

$$p_{i,j}^{B}=\frac{p_{i,j}}{P_B}=\frac{c_{i,j}}{\sum_{i=0}^{k}\sum_{j=k+1}^{L-1}c_{i,j}},\quad 0\le i\le k,\ k+1\le j\le L-1 \tag{25}$$

$$p_{i,j}^{C}=\frac{p_{i,j}}{P_C}=\frac{c_{i,j}}{\sum_{i=k+1}^{L-1}\sum_{j=k+1}^{L-1}c_{i,j}},\quad k+1\le i,j\le L-1 \tag{26}$$

$$p_{i,j}^{D}=\frac{p_{i,j}}{P_D}=\frac{c_{i,j}}{\sum_{i=k+1}^{L-1}\sum_{j=0}^{k}c_{i,j}},\quad k+1\le i\le L-1,0\le j\le k \tag{27}$$

B. Local and Joint Entropies

The second-order local entropy $H_{local}^{(2)}(k)$ and joint entropy $H_{joint}^{(2)}(k)$ are introduced by Pal and al.[11] to take advantage of spatial correlation in an image. They are defined as:

$$\begin{aligned}H_{local}^{(2)}(k)&=H_A^{(2)}(k)+H_C^{(2)}(k)\\&=-\frac{1}{2}\sum_{i=0}^{k}\sum_{j=0}^{k}p_{i,j}^{A}\log p_{i,j}^{A}-\frac{1}{2}\sum_{i=k+1}^{L-1}\sum_{j=k+1}^{L-1}p_{i,j}^{C}\log p_{i,j}^{C}\end{aligned} \tag{28}$$

$$\begin{aligned}H_{joint}^{(2)}(k)&=H_B^{(2)}(k)+H_D^{(2)}(k)\\&=-\frac{1}{2}\sum_{i=0}^{k}\sum_{j=k+1}^{L-1}p_{i,j}^{B}\log p_{i,j}^{B}-\frac{1}{2}\sum_{i=k+1}^{L-1}\sum_{j=0}^{k}p_{i,j}^{D}\log p_{i,j}^{D}\end{aligned} \tag{29}$$

$H_A^{(2)}(k)$ and $H_C^{(2)}(k)$ are the local entropies of objects and background respectively. $H_B^{(2)}(k)$ and $H_D^{(2)}(k)$ are the entropies of the edge information on transitions from objects to background and from background to objects, respectively. These two algorithms select thresholds that maximize the $H_{local}^{(2)}(k)$ and $H_{joint}^{(2)}(k)$ over *k*.

C. Relative Entropy

According to Kullback's relative entropy definition, Chang et al. [13] defined the relative entropy of the probability distributions $p_{i,j}$ and $p'_{i,j}$ as follows:

$$\begin{aligned}L(p;p')&=\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}p_{i,j}\log\frac{p_{i,j}}{p'_{i,j}}\\&=\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}p_{i,j}\log p_{i,j}-\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}p_{i,j}\log p'_{i,j}\end{aligned} \tag{30}$$

where $p_{i,j}$ and $p'_{i,j}$ are the transition probabilities from gray level *i* to gray level *j* of the original image and the bi-level image, respectively. The $p'_{i,j}$ can be defined as follows:

$$p'^{(A)}_{i,j}(k)=q_A(k)=\frac{P_A(k)}{(k+1)\times(k+1)},\ 0\le i,j\le k \tag{31}$$

$$p'^{(B)}_{i,j}=q_B(k)=\frac{P_B(k)}{(k+1)\times(L-k-1)},\ 0\le i\le k,\ k+1\le j\le L-1 \tag{32}$$

$$p'^{(C)}_{i,j}==q_C(k)=\frac{P_C(k)}{(L-k-1)\times(L-k-1)},\ k+1\le i,j\le L-1 \tag{33}$$

$$p'^{(D)}_{i,j}==q_D(k)=\frac{P_D(k)}{(L-k-1)\times(k+1)},\ k+1\le i\le L-1,0\le j\le k \tag{34}$$

The first term of (30) is independent of the threshold *k*. The second term of the equation is simplified as follows:

$$\begin{aligned}\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}p_{i,j}\log p'_{i,j}&=\sum_{A}p_{i,j}\log q_A(k)+\sum_{B}p_{i,j}\log q_B(k)\\&+\sum_{C}p_{i,j}\log q_C(k)+\sum_{D}p_{i,j}\log q_D(k)\\&=P_A(k)\log q_A(k)+P_B(k)\log q_B(k)\\&+P_C(k)\log q_C(k)+P_D(k)\log q_D(k)\end{aligned} \tag{35}$$

Minimizing *L(p;p')* is equivalent to maximizing the second term of the equation over *k*, to obtain the optimal threshold.

V. Locally Adapative Tresholding

A. Local Thresholding by Niblack Method

In some radiographic images, the background intensity is variable, and the overlapping between the two classes is therefore large, due to the weld thickness variations, the weak sizes of the defect and the geometrical considerations related to the used radiography technique. In such case, by a global thresholding, we do not obtain the desired results. That is why a local adaptive thresholding technique can be employed to overcome the problem. The method of Niblack is fast to implement and easy to apply. The main idea of Niblack's thresholding method [14] is to vary the threshold value over the input image, based on the local mean $\mu(x,y)$ and local standard deviation $\sigma(x,y)$. The threshold value at pixel (x,y) is computed by

$$T(x,y) = \mu(x,y) + k\sigma(x,y) \tag{36}$$

where k is a parameter which depends on image content.

The size of the neighborhood must be sufficiently small to preserve the local details but also, it must be enough large to remove the noise. In this method, the problem is the case of the presence of the light textures in the background, which are considered as object with small contrast.

B. Local Thresholding by Sauvola Method

To overcome the above mentioned problems, Sauvola [15] proposed a new improved formula to compute the threshold

$$T(x,y) = \mu(x,y)\left(1 - k\alpha\right) \tag{37}$$

where

$$\alpha = 1 - \sigma(x,y)/R \tag{38}$$

k: positive value parameter. R: dynamic range of variance.

The contribution of the standard deviation becomes adaptive. In this method, hypothesis on the gray levels of the object and the background are used to eliminate the noise produced by light textures of the background because μ reduces the threshold value in the light background regions.

VI. Thresholding Performance Criteria

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 [16]. 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 the thresholding methods [5], we have used the following four performance criteria: misclassification error (ME), region non uniformity (NU), relative foreground area error (RAE) and shape measure (SM). 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) [17] reflects the percentage of background pixels wrongly assigned to foreground, and conversely, foreground pixels wrongly assigned to background. It can be simply expressed as:

$$ME = 1 - \frac{\left|B_O \cap B_k\right| + \left|F_O \cap F_k\right|}{\left|B_O\right| + \left|F_O\right|} \tag{39}$$

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.

B. Region Non-Uniformity

The region non uniformity criterion [18] [16] is not based on ground truth data, but judges the intrinsic quality of the segmented areas. This measure is defined as

$$NU = \frac{\left|F_k\right|}{\left|F_k + B_k\right|}\frac{\sigma_0^2}{\sigma_T^2} \tag{40}$$

where σ_T^2 represents the variance of the whole image, and σ_0^2 represents the foreground variance.

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 [16] to reflect the feature measurement accuracy. This measure is modified for the area feature A as follows:

$$\text{RAE} = \begin{cases} \dfrac{A_O - A_k}{A_O} & \text{if} \quad A_k < A_O \\ \dfrac{A_k - A_O}{A_k} & \text{if} \quad A_k \geq A_O \end{cases} \tag{41}$$

where A_0 and A_k are the foreground areas in the reference image and the thresholded image, respectively.

D. Goodness based on Region Shape

Not only the gray level, but also the form of a segmented region can be taken into account to design goodness for satisfying the human intuition on an "ideal" segmentation. So, to account for shape inaccuracies, Sahoo et al. [19] proposed a shape measure (SM) for threshold evaluation. They assume a (3×3) neighborhood centered around the test pixel (x,y) with gray value $f(x,y)$, and define a general gradient $\Delta(x,y)$, as the root-mean-square vector sum of the grey level difference in all four directions. The original version of the shape measure is slightly modified in order to adjust this performance measure so that its scores vary from 0 for a correct segmentation to 1 for an erroneous case. The shape measure SM at threshold k is then given by:

$$SM = 1 - \frac{\sum_{x,y} Sgn\left(f(x,y) - \overline{f_N(x,y)}\right)\Delta(x,y)Sgn\left(f(x,y) - k\right)}{C} \quad (42)$$

where, $Sgn(x)$ = -1 or 1 depending on whether or not x is negative, $\overline{f_N(x,y)}$ is the average grey level of the neighborhood $N(x,y)$ and C is a normalization factor based on the region area and its gray level range.

The computation of the generalized gradient value $\Delta(x,y)$ of the pixel (x,y) is carried out using the formula

$$\Delta(x,y) = \left[\sum_{i=1}^{4} D_i^2 + \sqrt{2}D_1\left(D_3 + D_4\right) - \sqrt{2}D_2\left(D_3 - D_4\right)\right]^{1/2} \quad (43)$$

where
$$\begin{cases} D_1 = f(x+1,y) - f(x-1,y), \\ D_2 = f(x,y-1) - f(x,y+1), \\ D_3 = f(x+1,y+1) - f(x-1,y-1), \\ D_4 = f(x+1,y-1) - f(x-1,y+1). \end{cases}$$

The shape measure expects a threshold such that all the points with positive gradient be above the threshold, and the points with negative gradient be below the threshold. The term shape measure seems to be a misnomer, because it depends mainly on the assumption that regions are uniform and the grey level gradients in the image correspond to the transitions between the two regions to be separated.

For the locally adaptive thresholding methods, we have adapted the shape measure so that it holds in account the spatial dependant threshold $T(x,y)$ to give us the following version of the criterion:

$$SM = 1 - \frac{\sum_{x,y} Sgn\left(f(x,y) - \overline{f_N(x,y)}\right)\Delta(x,y)Sgn\left(f(x,y) - T(x,y)\right)}{C} \quad (44)$$

E. 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, RAE and SM criteria. In other words, given a thresholding algorithm, for each image the average of ME, NU, RAE and SM 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:

$$S(i) = \left(ME(i) + NU(i) + RAE(i) + SM(i)\right)/4 \quad (45)$$

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

$$S_t = \left(\sum_{i=1}^{4} S(i)\right) \Big/ n \quad (46)$$

with n the number of the tested images.

VII. Experimental Results and Discussion

In order to show the effectiveness of the different 1-D and 2-D histogram-based thresholding methods in one part and the locally adaptive thresholding methods in the other hand on real data, a set of 11 radiographic images representing weld defects such as lack of penetration, transversal and longitudinal cracks, undercut, solid inclusions and porosities was used. The weld defect images and their corresponding ground truths are shown in Figures 2, 3 and 4. 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 [20] such as uneven background illumination. Nevertheless, an appropriate contrast enhancement technique can contribute in the improvement of the thresholding quality. The binary images obtained by the nine studied thresholding methods are shown in Figures 2, 3 and 4. The performance measures of the proposed methods and their ranking (between hooks) are reported in the table I. For the second order entropy methods, after a series of test, the values of $d = 5$ and $\theta = 45°$ are taken because they prove to be the most adequate for the whole of algorithms and images. For the Niblack and Sauvola methods, the values of $W=13$, $k(Nib.)=-0.2$, $k(Sauv.)=0.5$ and $R=128$ are selected. This choice was made in an empirical way, taking in account the dilemma between robustness (non sensitiveness to noise) and precision (space definition of the segmented areas).

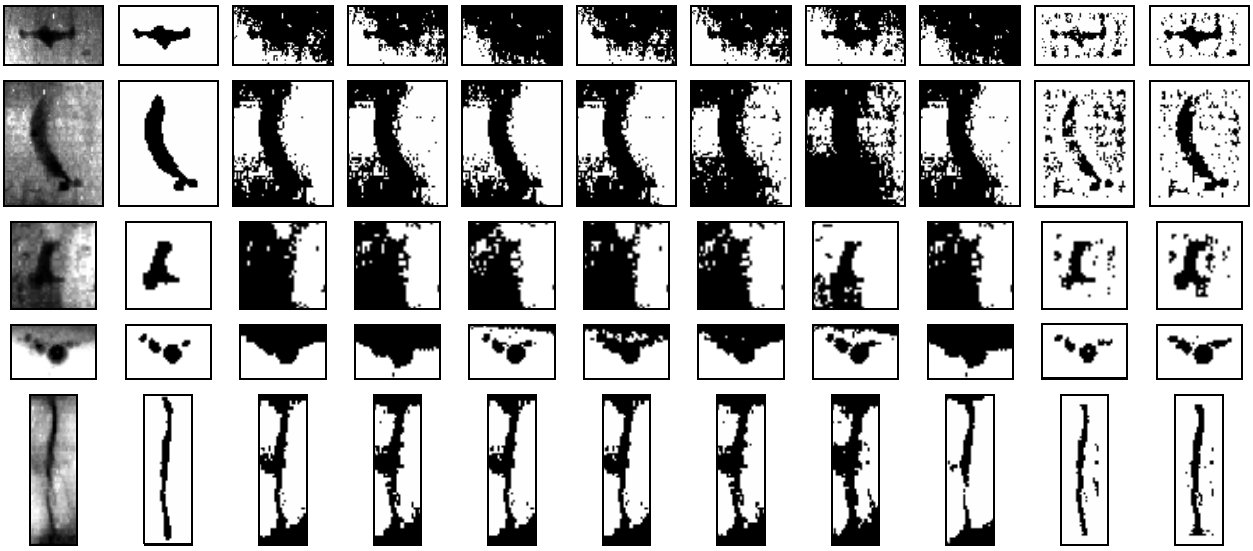

Fig. 2 (from top to bottom : Images 1 to 5) (from left to right: Original weld defect image, Ground truth, Thresholding by Otsu, Kittler, Kapur, Tsai, Local entropy, Joint entropy, Relative entropy, Niblack and Sauvola)

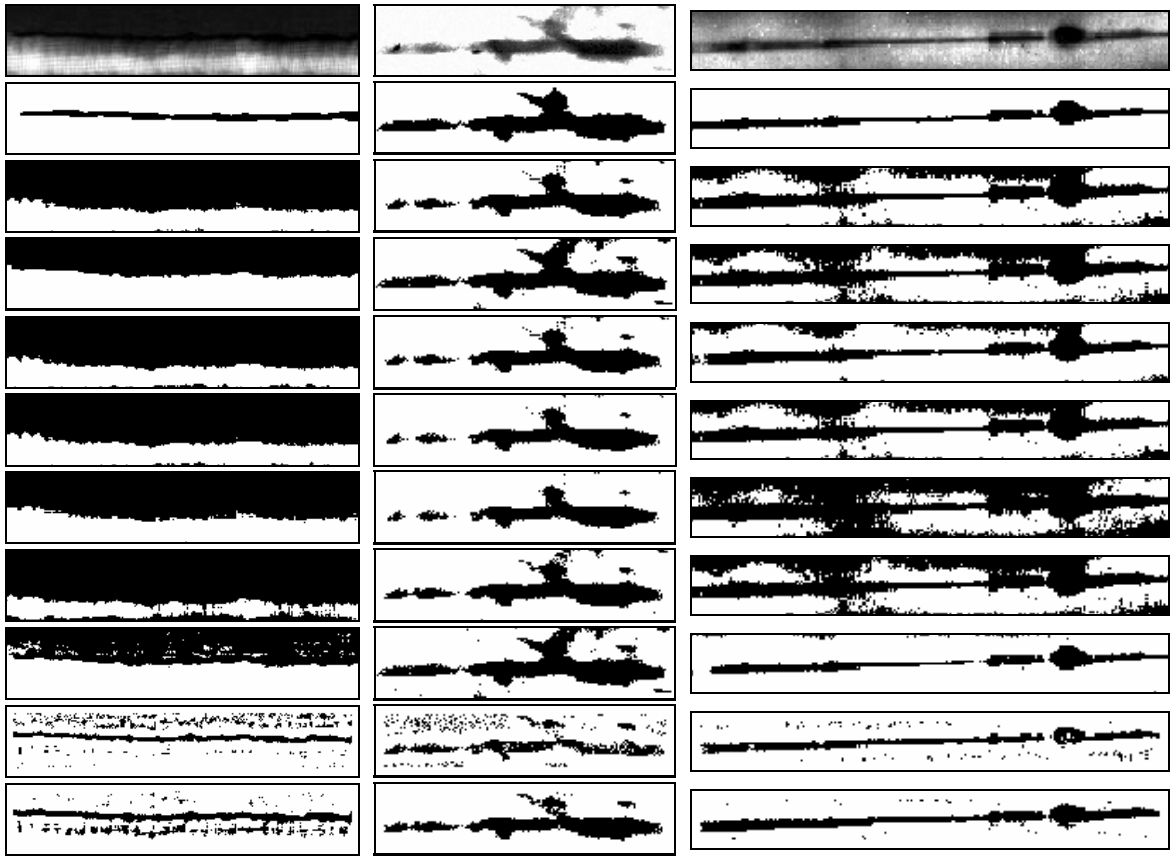

Fig. 3 (from left to right : Images 6 to 8) (from top to bottom: Original weld defect image, Ground truth, Thresholding by Otsu, Kittler, Kapur, Tsai, Local entropy, Joint entropy, Relative entropy, Niblack and Sauvola)

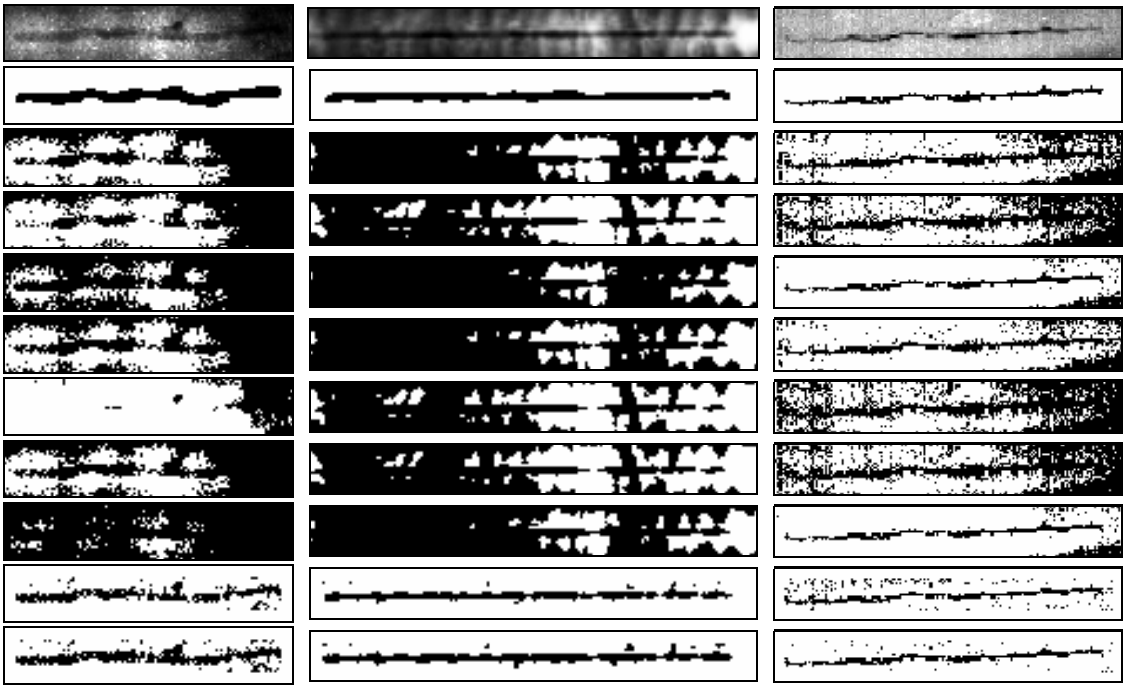

Fig. 4 (from left to right : Images 9 to 11) (from top to bottom: Original weld defect image, Ground truth, Thresholding by Otsu, Kittler, Kapur, Tsai, Local entropy, Joint entropy, Relative entropy, Niblack and Sauvola)

For 1-D Histogram based approach; the techniques just consider the overall gray level statistics of the input image and don't involve the relationship between the pixels. Thus they have the advantage of faster processing speed and straightforward great result for simple images. So when the input images have clearly classified histogram with two peaks distribution, it's suitable to use this kind of approach.

For 2-D Histogram based approach, like local, joint and relative entropy methods, the relationship between pixels is considered. So when the input image presents some complicate relationship or overlap between foreground and background, the 2-D Histogram approach can give us a better result compare to the 1-D Histogram approach. But we can't guarantee the 2-D Histogram always better than 1-D Histogram since the result really depends on the particular input image and the desired foreground.

In the majority of the proposed weld defect images, the background intensity is variable, and the overlapping between the two classes is therefore large, due to the weld thickness variations, the weak sizes of the defect and the geometrical considerations related to the used radiography technique. In such case, by a global thresholding, we do not obtain the desired results. That is why a local adaptive thresholding technique can be used to overcome the problem.

TABLE I
THRESHOLDING EVALUATION RESULTS FOR WELD DEFECT IMAGES

		Im.1	*Im.2*	*Im.3*	*Im.4*	*Im.5*	*Im.6*	*Im.7*	*Im.8*	*Im.9*	*Im.10*	*Im.11*	S_t
Otsu	S	0.5863 [7]	0.4925 [4]	0.5575 [9]	0.4443 [7]	0.4560 [4]	0.6142 [5]	0.1988 [4]	0.4345 [6]	0.5025 [5]	0.5694 [6]	0.4705 [6]	0.4842 [6]
Kittler	S	0.5309 [4]	0.5157 [7]	0.5286 [6]	0.5809[8]	0.5102 [7]	0.5585 [4]	0.2307 [6]	0.4913 [7]	0.4875 [4]	0.4843 [3]	0.5560 [7]	0.4977 [9]
Kapur	S	0.6768 [9]	0.4750 [3]	0.4985 [4]	0.2723 [3]	0.4667 [5]	0.6244 [8]	0.1984 [3]	0.3578 [4]	0.5870 [8]	0.6432 [9]	0.2937 [3]	0.4631 [3]
Tsai	S	0.5600 [6]	0.4997 [5]	0.5409 [8]	0.3843 [5]	0.4667 [5]	0.6189 [7]	0.2270 [5]	0.4343 [5]	0.5144 [6]	0.5767 [7]	0.4104 [5]	0.4758 [4]
Loc. Ent.	S	0.5503 [5]	0.5529 [8]	0.5203 [5]	0.4253 [6]	0.5137 [8]	0.6031 [5]	0.2413 [8]	0.6146 [9]	0.3427 [3]	0.4913 [4]	0.5869 [9]	0.4948 [8]
Joi. Ent.	S	0.5037 [3]	0.6311 [9]	0.3719 [3]	0.3179 [4]	0.5630 [9]	0.6989 [9]	0.1745 [1]	0.4979 [8]	0.5151 [7]	0.5061 [5]	0.5630 [8]	0.4857 [7]
Rel. Ent.	S	0.6725 [8]	0.5151 [6]	0.5286 [6]	0.5976 [9]	0.3617 [3]	0.5471 [3]	0.2333 [7]	0.2057 [2]	0.7229 [9]	0.6045 [8]	0.2937 [3]	0.4802 [5]
Nibl.	S	0.3260 [2]	0.3259 [2]	0.2052 [1]	0.1617 [1]	0.1712 [1]	0.3307 [2]	0.3517 [9]	0.1911 [1]	0.2533 [2]	0.1384 [2]	0.2417 [2]	0.2452 [2]
Sauv.	S	0.2993 [1]	0.2820 [1]	0.2366 [2]	0.1700 [2]	0.2341 [2]	0.3148 [1]	0.1930 [2]	0.2144 [3]	0.1845 [1]	0.1066 [1]	0.1970 [1]	0.2211 [1]

By examining the thresholding scores we can deduce that for the overall images, except image 8, the best result were provided by the locally adaptive methods, especially by the Sauvola thresholding method, nevertheless the post processing is needed, essentially for the image 1, 2 and 3 to remove the

noise and isolated spots, whereas the better result for Image 8 was provided by the relative entropy method.

So, for the selected radiographic images generally, the global 1-D and 2-D histogram-based methods results are not satisfactory. This substandard performance can be explained by the fact that these images present non uniform intensity for the background which confounds in some areas with the defect region. For example, for all the global methods, the binarized version of the weld defect (external undercut) presented in image 6 is totally drowned in the background. This can affect the results of interpretation dangerously. The same remark can be done for images 1, 3, 4 and 10 where the weld defect aspect cannot be arisen, except for Kapur and joint entropy methods in the case of image 4. Still for overall images, the Kapur thresholding method outperforms all the other methods in the category of 1-D histogram-based approach whereas the Relative Entropy method is the best in the category of 1-D histogram-based approach. So, except for the locally adaptive methods 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. It is noted that the 2-D Histogram-based algorithms are the slowest methods in terms of calculation time, and the 1-D histogram-based methods are the fasters since they have much less calculations to process.

VIII. Conclusion

In this study we have investigated experimentally the effectiveness of 1-D and 2-D histogram-based and the locally adaptive thresholding methods through radiographic images of welds. To evaluate the quality of these thresholding techniques, a performance criterion is measured on the resulted binary images. Compare all of these nine different methods, for well contrasted images, the Kapur method is the best for the 1-D histogram-based methods. The relative entropy is slightly more powerful than the other methods in the 2-D histogram-based approach. According to the result scores on the overall weld defect images, the locally adaptive methods, especially, the Sauvola method, prove to be the stronger thresholding tool for this kind of images.

To summarize the main results presented in this paper, we note that generally, the global thresholding methods gives good results for well contrasted weld defect radiographic images. In the case of images with non uniform background intensity, the methods of Niblack and Sauvola are recommended. Nevertheless, in the Niblack's method, the problem lies in the light textures of the background, which are assimilated to objects with low contrast. To overcome this problem, the method of Sauvola can be applied.

In general, 1-D Histogram based algorithm is much faster than 2-D Histogram based algorithm and the locally adaptive methods since much less calculations are needed.

References

[1] J. H. Rogerson, "Defects in welds – Their prevention and their significance". *Applied science publishers*, 1983.
[2] De Carvalho & al., "Evaluation of the relevant features of welding defects in radiographic inspection," *Materials Research*, vol. 6, n° 3, pp. 427-432, 2003.
[3] Ch. Schwartz, "Automatic Evaluation of Welded Joints Using Image Processing on Radiographs," Conference Proceedings *American Institute of Physics,* vol 657(1) pp. 689-694, 2003.
[4] L. Soler, G. Malandrin, H. Delinguette, "Segmentation automatique: Application aux angioscanner 3D," *Revue de Traitement de Signal*, vol. 15, 1998.
[5] M. Sezgin, B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *Journal of Electronic imaging* 13(1), Jan. 2004, pp. 146–165.
[6] S.U. Lee, S.Y. Chung, 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.
[7] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. on Systems, Man, and Cybern.* vol. SMC-9, 1979, pp. 62-66.
[8] J. Kittler and J. Illingworth, "Minimum Error Thresholding," *Pattern Recognition,* vol. 19, no.1, 1986.
[9] 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.
[10] W-H. Tsai, "Moment-Preserving Thresholding: A New Approach," *Computer Vision, Graphics, and Image Processing*, vol. 29, 1985, pp. 377-393.
[11] N. R. Pal and S. K. Pal, "Entropy thresholding," *Signal Processing*, vol. 16, 1989, pp. 97–108.
[12] S.S. Lee, S.J. Horng and H.R. Tsai, "Entropy Thresholding and Its Parallel Algorithm on the Reconfigurable Array of Processors with Wider Bus networks," *IEEE Trans. Image Processing*, vol. 8 no. 9, Sept. 1999, pp.1229-1242.
[13] C.I. Chang, K. Chen, J. Wang and M.L.G. Althouse, "A Relative entropy-based approach to image thresholding, *Pattern Recognition*, vol. 27. no.9, 1994, pp. 1275-1289.
[14] W. Niblack, *Introduction to digital image processing*. Prentice Hall, 1986.
[15] J. Sauvola, M. Pietikakinen , "Adaptive document image binari-zation". *Pattern Recognition*, 33, 2000 pp. 225-236.
[16] Y. J. Zhang, "A survey on evaluation methods for image segmentation," *Pattern Recognition*, vol. 29 no.8, 1996.
[17] W. A. Yasnoff, J. K. Mui, and J. W. Bacus, "Error measures for scene segmentation," *Pattern Recognition,* vol. 9, 1977, pp. 217–231.
[18] M. D. Levine and A. M. Nazif, "Dynamic measurement of computer generated image segmentations," *IEEE Trans. PAMI*, vol. 7, 1985, pp. 155–164.
[19] P. K. Sahoo, S. Soltani, A. K. C. Wong, "A survey of thresholding techniques". Computer Vision, Graphics, and Image Processing, vol. 41, 1988, pp. 233-260.
[20] N. Nacereddine, "Weld defect detection and recognition in industrial radiography based image analysis and neuronal classifiers," *4th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurised Components*, London, 6-8 Dec. 2004