

An object detection and recognition system for weld bead extraction from digital radiographs

Marcelo Kleber Felisberto, Heitor Silvério Lopes, Tania Mezzadri Centeno *,
Lúcia Valéria Ramos de Arruda

CPGEI/CEFET-PR, Av. Sete de Setembro, 3165 CEP: 80230-000 Curitiba-PR, Brazil

Received 2 August 2005; accepted 14 February 2006

Available online 31 March 2006

Abstract

With base in object detection and recognition techniques, we developed and implemented a new methodology to perform the first head-function of a weld quality interpretation system: the weld bead extraction from a digital radiograph. The proposed methodology uses a genetic algorithm to manage the search for suitable parameters values (position, width, length, and angle) that best defines a window, in the radiographic image, matching with the model image of a weld bead sample. The search results are verified in a classification process that recognize true detections using image matching parameters also proposed in this work. To test the proposed methodology, two groups of images were used; one consisting of 110 radiographs from pipelines welded joints and the other containing 6 images with different numbers of radiographs per image. The tests results showed that, besides automatically check the number of weld beads per image, the proposed methodology is also able to supply the respective position, width, length, and angle of each weld bead, with an accurate rate of 94.4%. As a result, the detected weld beads are correctly extracted from the original image and made available to be inspected through others algorithms for failure detection and classification.

© 2006 Published by Elsevier Inc.

Keywords: Object detection; Image matching; Genetic algorithms; Image segmentation; Radiographic weld inspection

1. Introduction

Radiographic inspection has been one of the most widely used non-destructive techniques for detection of internal defects in welded structures in the industry [14]. Currently, the analysis of the weld joint radiographs is performed by qualified inspectors, who make decisions based in their own experience and visual accuracy. Consequently, results are often subjective and many efforts have been done towards the design and construction of computer vision systems, aiming at supporting the weld joint radiographs interpretation and, therefore, improving robustness, accuracy, and speed of the inspection process [15].

According to Liao [13], a weld radiographs interpretation system generally has three major functions: (1) segmentation of the welds from the background, (2) detection of welding defects in the weld, and (3) classification of the defects' types. Liao and Ni [10] also emphasize that, since only items within a weld are of interest, it is preferred to extract the weld from each image before applying defect detection algorithms. Therefore, this article presents a new approach, based on object detection and recognition techniques, for the first step of this image analysis process, i.e., segmentation of the welds from the image background.

Concerned with the weld bead segmentation, one of the first related works to be quoted is the approach from Lawson and Parker [9] that uses artificial neural networks. Basically, they used a multilayer-perceptron (MLP) neural network (NN), trained by error backpropagation, in order to recognize weld bead image pixels from the background image in a digitized radiograph. However, with their NN

* Corresponding author. Fax: +55 41 310 4683.

E-mail addresses: mkf@cpgei.cefetpr.br (M.K. Felisberto), hslopes@cpgei.cefetpr.br (H.S. Lopes), mezzadri@cpgei.cefetpr.br (T.M. Centeno), arruda@cpgei.cefetpr.br (L.V.R. de Arruda).

training, they chose some pixels' characteristics that are closely dependent to the pixel position in the image. As a consequence, the system effectiveness depends on the previous knowledge about the width, position and angle of the weld.

Another approach, concerned with weld bead segmentation, was introduced by Liao and Ni [10]. Their method was based on the observation that if the joint was welded with excess of metal its transversal profile of intensities (gray levels) is more Gaussian-like than other objects in the image. Based on the peaks and valleys position along each line image profile, line intercepted objects are firstly detected. Then, each object profile is compared to a Gaussian curve in terms of the mean square error (MSE). In further works, the MSE, as well as other object profiles characteristics (height, width, and position), are used for pixels classification into two classes: weld and non-weld, using fuzzy-based classifiers [12] and MLP neural networks [11]. However, depending on the weld angle and position, a line image cannot contain the weld transversal but longitudinal profile, which is not Gaussian-like. Therefore, a previous knowledge about the weld angle is required in this case. In addition, due to the fact that some shaved/repaired welds usually have non-Gaussian characteristics, this implies that another methodology is also needed in these cases [10].

In our previous work [4], we introduced an object detection system, based on evolutionary computation and image matching techniques, to find specific chess pieces in images with different scales, positions, angles, and number of objects. Using the same principles, we adapted and improved the method to be used in the weld bead segmentation problem. As a result, the system was enabled to detect the position, width, length, and angle of how many weld beads are there in the image with an accuracy rate of 94.4%.

In the sequence, this paper is organized as follows: problem characterization in Section 2, methodology in Section 3, tests and results in Section 4, and final conclusions in Section 5.

2. Problem characterization

All the digital radiographic images that were acquired (111 in total) came from carbon-steel pipes with 6 or 8 in. nominal diameter and a wall thickness between 6.4 and 12.7 mm. The pipes were welded by gas tungsten arc welding (GTAW) process, using carbon-steel filler metal. Instead of traditional X-ray films, we used “photostimulable” phosphor plates (manufactured by Gendex®) with 150 mm × 300 mm. After the X-ray radiation exposition, each phosphor plate is read by a laser scanner device (DenOptix®, manufactured by Gendex®) and converted into a digital image with 256 grey levels at 150 dpi (dots per inch). It is important to emphasize that this resolution is not proper for analysis of defects and it was used just for the weld bead detection. The radiographic test was

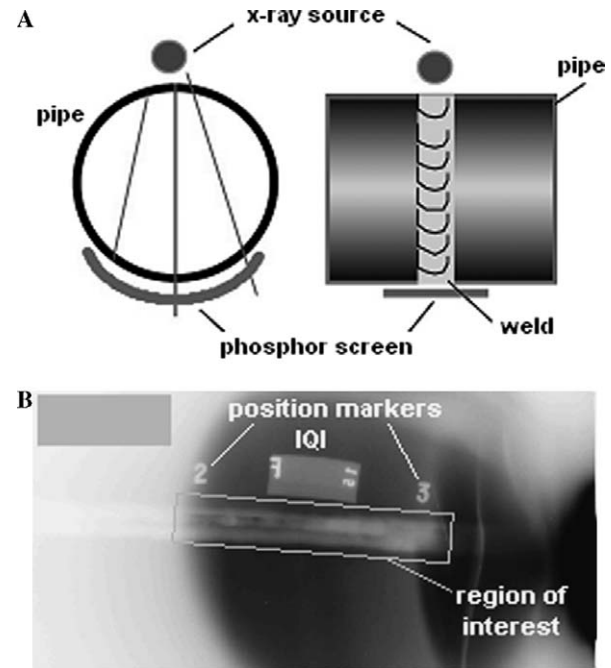


Fig. 1. (A) Double-wall radiographic technique with single-wall radiographic viewing. (B) Sampled radiograph from a pipe weld joint.

conducted according to ASME V [1], using double-wall radiographic exposing technique with single-wall radiographic viewing, as shown in Fig. 1A. Evidently, such technique has to be repeated in different angles to embrace the pipe weld joint and cover the whole welded bead. Fig. 1B shows a sampled image, which were acquired through this technique.

In our approach, the main objective is to automatically detect the region of interest, which is the weld bead stretch indicated by the two position markers, as shown in Fig. 1B. Some objects, as the position markers and image quality indicator (IQI), increase the complexity of the problem. They make the weld image segmentation by a simple gray level threshold impossible. In addition, many parameters (width, length, angle, position, and contrast of the weld bead) can vary from one image to another, making the task of weld bead detection even more difficult.

3. Methodology

The object detection and recognition system, which the block diagram is shown in Fig. 2, was adapted for our problem. This system was first idealized by Centeno et al. [4] to find chess pieces in images. Actually, each block function is explained as follows:

- The *model image* (an object sample image) is fed to the model construction block. This block processes it in order to construct the model representation (described in Section 3.1).

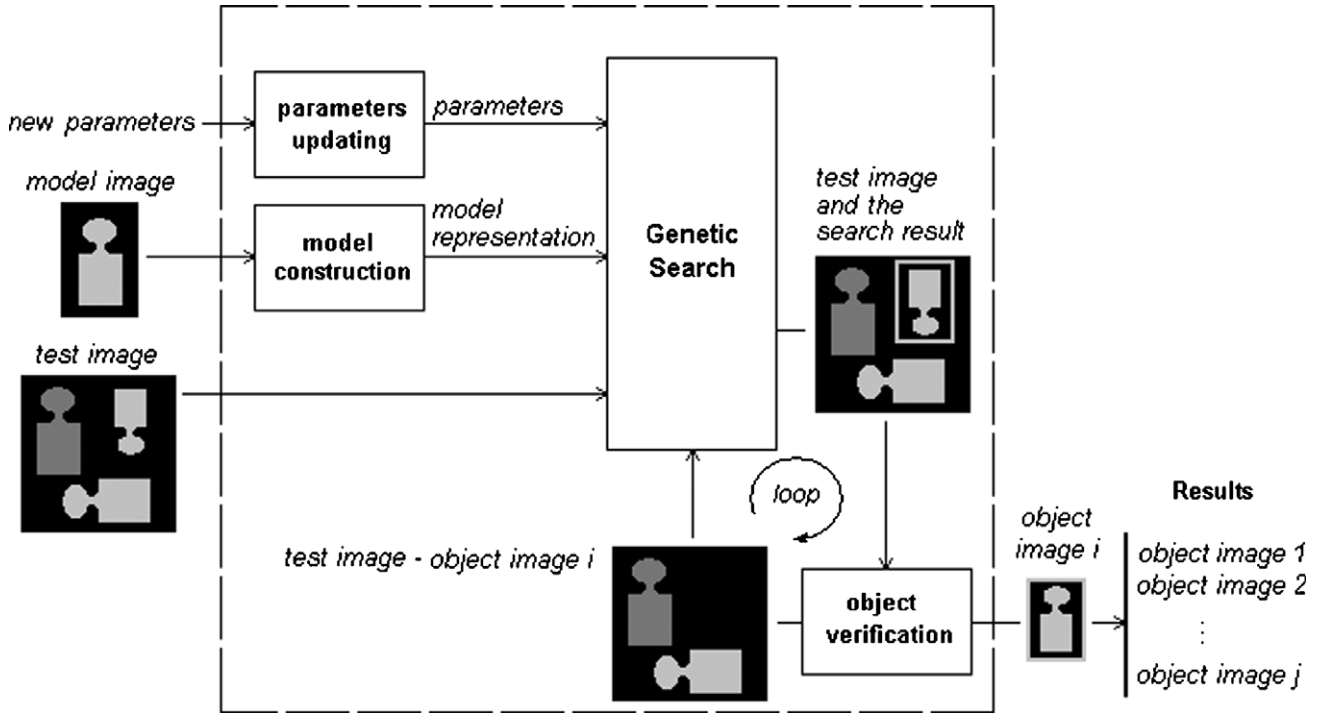


Fig. 2. The block diagram showing how the information flows through the components (blocks) of the object detection and recognition system [4].

- The *test image* (the image to be analyzed) is fed to the genetic search block. This block uses a genetic algorithm (GA) in order to find the region of the *test image* that better matches the *image model* (Sections 3.2 and 3.3).
- The object verification block evaluates the GA output (*object image i*). Each time *object image i* is accepted, it is extracted from the test image, *i* is updated to (*i* + 1) and the remaining image (*test image*—*object image i*) is feed-backed to the GA for another search. As a consequence, when the current *object image i* is not accepted, the search stops. Besides, (*i* − 1) gives the number of detected objects. The loop between the blocks genetic search and object verification allows the system to find further copies of the same pattern in the *test image*.
- The parameters updating block is just for the system parameters adjustment.

3.1. Model construction

The *image model* for this study was the weld bead radiograph from Fig. 3A. As in Centeno et al. [4], the *image model* is represented by a $(n \times n)$ matrix, called reference matrix (*Mref*), with $n = 14$. The procedure to generate *Mref* is described as follows:

- The *model image* is divided by n horizontal lines regularly spaced by distance dy (where $dy = \text{number of image lines}/(n + 1)$, in pixels).
- The *model image* is divided by n vertical lines regularly spaced by distance dx (where $dx = \text{number of image columns}/(n + 1)$, in pixels).

- The intersection points (see Fig. 3A) are named ‘reference points’ and are denoted by P_{ij} , where $i, j = 1, \dots, n$.
- Point P_{cc} , with $c = (1 + n/2)$, is named the ‘central reference point.’
- Function $f(P_{ij})$ assigns to each reference point the value corresponding to the mean gray level of the pixels in the P_{ij} neighborhood, as defined by the limited region shown in Fig. 3B.
- All the $f(P_{ij})$ values, normalized in the integer range [0..255], are represented in the $(n \times n)$ reference matrix (*Mref*), as Eq. (1):

$$M_{\text{ref}} = \begin{bmatrix} f(P_{11}) & \dots & f(P_{1n}) \\ \dots & f(P_{ij}) & \dots \\ f(P_{n1}) & \dots & f(P_{nn}) \end{bmatrix}. \quad (1)$$

Besides *Mref*, the following parameters were also studied in this work:

- Mref* derivative in the i direction

$$\delta M_{\text{ref}} / \delta i(i, j) = M_{\text{ref}}(i, j+1) - M_{\text{ref}}(i, j), \text{ for } i = 1, \dots, n \text{ and } j = 1, \dots, n-1 \quad (2)$$

- Mref* derivative in the j direction

$$\delta M_{\text{ref}} / \delta j(i, j) = M_{\text{ref}}(i+1, j) - M_{\text{ref}}(i, j), \text{ for } i = 1, \dots, n-1 \text{ and } j = 1, \dots, n. \quad (3)$$



Fig. 3. (A) Model image divided by $n = 14$ lines and columns, with the central point of reference P_{cc} . (B) The P_{ij} neighborhood for a single (i,j) point.

(c) M_{ref} horizontal projection

$$Ph(M_{ref})(i) = \sum_{j=0}^{n-1} M_{ref}(i, j) \quad \text{for } i = 1, \dots, n. \quad (4)$$

(d) M_{ref} vertical projection

$$Pv(M_{ref})(j) = \sum_{i=0}^{n-1} M_{ref}(i, j) \quad \text{for } j = 1, \dots, n. \quad (5)$$

The sample, in Fig. 4, shows a (3×3) M_{ref} and its derivatives ($\delta M_{ref}/\delta i$ and $\delta M_{ref}/\delta j$) and corresponding projections ($Ph(M_{ref})$ and $Pv(M_{ref})$).

3.2. Image matching

Supposing that *image k* is a rectangular region (window) extracted from the *test image*, Centeno et al. [4] use a function $F(\text{image } k)$ to compute the similarity between the *image k* and the *model image*:

$$F_G(\text{image}_k) = \frac{w_F \cdot F(\text{image}_k) + w_{F_{di}} \cdot F_{di}(\text{image}_k) + w_{F_{dj}} \cdot F_{dj}(\text{image}_k) + w_{F_{ph}} \cdot F_{ph}(\text{image}_k) + w_{F_{pv}} \cdot F_{pv}(\text{image}_k)}{w_F + w_{F_{di}} + w_{F_{dj}} + w_{F_{ph}} + w_{F_{pv}}}, \quad (13)$$

$$F(\text{image } k) = S(M_{ref}_k, M_{ref}). \quad (6)$$

The similarity value $S(M_{ref}_k, M_{ref})$ is computed for the reference matrices: M_{ref}_k (for the *image k*) and M_{ref} (for the *model image*) in terms of the absolute error sum, using Eqs. (7) and (8). Indeed, these equations give the similarity between two $(n \times m)$ matrices (M_1 and M_2)

$$S(M_1, M_2) = 1 - E/E_{\max}, \quad (7)$$

where

$$E = \sum_{i=1}^n \sum_{j=1}^m |M_1(i, j) - M_2(i, j)| \quad (8)$$

and E_{\max} is the maximum value for E .

Besides F , four other image-matching parameters (F_{di} , F_{dj} , F_{ph} , F_{pv}) were studied in this work. The parameters F_{di} and F_{dj} are related to the reference matrices derivatives, according to Eqs. (9) and (10). Similarly, the parameters F_{ph} and F_{pv} are related to the horizontal and vertical projections, according to Eqs. (11) and (12).

	M_{ref}	$Ph(M_{ref})$	$\delta M_{ref}/\delta j$
	$\begin{bmatrix} 50 & 30 & 150 \\ 100 & 150 & 230 \\ 200 & 100 & 80 \end{bmatrix}$	$\begin{bmatrix} 230 \\ 480 \\ 380 \end{bmatrix}$	$\begin{bmatrix} -20 & 120 \\ 50 & 80 \\ -100 & -20 \end{bmatrix}$
$Pv(M_{ref})$	$\begin{bmatrix} 350 & 280 & 460 \end{bmatrix}$		
$\delta M_{ref}/\delta i$	$\begin{bmatrix} 50 & 120 & 80 \\ 100 & -50 & -150 \end{bmatrix}$		

Fig. 4. A matrix (3×3) M_{ref} , derivatives $\delta M_{ref}/\delta i$ and $\delta M_{ref}/\delta j$, and projections $Ph(M_{ref})$ and $Pv(M_{ref})$.

$$F_{di}(\text{image } k) = S(\delta M_{ref}_k/\delta i, \delta M_{ref}/\delta i), \quad (9)$$

$$F_{dj}(\text{image } k) = S(\delta M_{ref}_k/\delta j, \delta M_{ref}/\delta j), \quad (10)$$

$$F_{ph}(\text{image } k) = S(Ph(M_{ref}_k), Ph(M_{ref})), \quad (11)$$

$$F_{pv}(\text{image } k) = S(Pv(M_{ref}_k), Pv(M_{ref})). \quad (12)$$

For convenience, instead of five matching parameters (F , F_{di} , F_{dj} , F_{ph} , F_{pv}), the weighted sum $F_G(\text{image } k)$ from Eq. (13), also called global similarity value, is used to compare one image (*image_k*) to the other (*model image*).

The variables w_F , $w_{F_{di}}$, $w_{F_{dj}}$, $w_{F_{ph}}$, and $w_{F_{pv}}$ are binary weights (0 or 1) for the matching parameters in the sum. In the case that any weight is turned to 0, the respective matching parameter is just cut out of the F_G formula. It is also important to point out that F_G varies in the range [0.00 1.00] and, actually, values closer to 1.00 are expected for the most similar images.

As shown in Fig. 5, any *image k* (Fig. 5B) can be defined within the *test image* (Fig. 5A) by an enclosing rectangle (image window). The position, scales, and angle of the enclosing rectangle (for *image k*) are defined by vector I_k , as follows:

$$I_k = (x_k, y_k, sh_k, sv_k, \theta_k), \quad (14)$$

where x_k and y_k are the coordinates of the central reference point (P_{cc}), sh_k gives the horizontal scale, sv_k gives the vertical scale and θ_k gives its rotation angle. Using these variables, *image k* can be extracted from the *test image* and, using the global similarity $F_G(\text{image } k)$, the *image k* can also be compared to the *model image* (Fig. 5C).

Unlike our previous work [4], independent scales (sh_k and sv_k) were used for the horizontal (direction of the

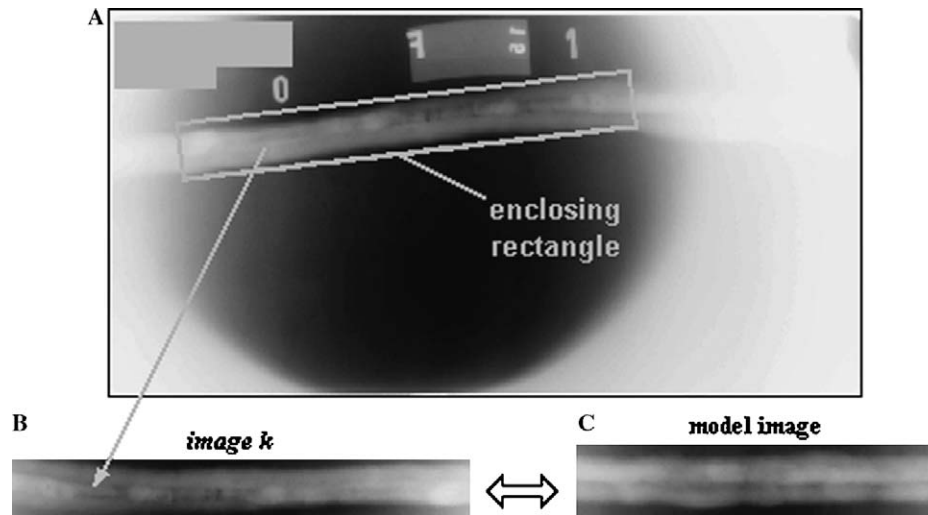


Fig. 5. (A) The *test image*—a typical weld joint radiograph. (B) The image k —a rectangular region extracted from the *test image*. (C) The *model image*.

weld bead length) and vertical (direction of the weld bead width) rectangle dimensions. However, similarly to our previous work, a genetic algorithm (GA) was also implemented to manage the search for the optimum I_k .

3.3. Genetic algorithms

GAs are well-known search and optimization techniques based on the mechanics of natural genetics. According to Simunic and Loncaric [16], GAs are very appropriate for optimization in large search spaces, where exhaustive search procedures are not feasible. Recent works [3,5,2] has also applied evolutionary computation techniques, such as GA, in order to improve image matching and segmentation techniques for object detection and recognition, with many advantages.

GAs uses population-based strategies where individuals are usually represented as strings of bits that, when suitably decoded, correspond to a feasible solution for the problem. Each individual can be evaluated by a fitness function whose value represents how good the current solution for the problem is. Starting from an initial population of individuals randomly generated, subsequent generations are evolved based on the natural selection principle and genetic operators like crossover and mutation. The Darwinian principle of the survival of the fittest gives more chance to the best individuals to propagate their own genetic material to the next generations. Consequently, it is expected that the most feasible solutions for the problem will be represented by individuals from the last generation [6].

In the next subsections it will be explained how GAs were used in our problem. This includes describing how the individuals are encoded (Section 3.3.1), how they are evaluated (3.3.2) and which genetic operators and parameters were used in our application (3.3.3).

3.3.1. Individual encoding

An individual I_k is defined by Eq. (14). Therefore, a general solution for the problem will have the following variables: x , y , sh , sv , and θ . Using a binary code (string) with 11 bits long to represent each variable, it is possible to express 2048 values within the feasible variable range. By using the range specified in Table 1 for the angle θ , for example, it shall vary in steps of 0.003 radian from 0 until 2π .

It is important to point out that the range in Table 1 for variable sv was set based on the estimated widths for the thickest and sharpest weld bead samples. Similarly, sh range was set based on the estimated lengths for the largest and shortest weld bead samples. The ranges for x and y , however, were defined in terms of the *test image* dimensions to assure that the central point of reference (P_{cc}) will always be a point within the *test image*. Since the images used in this work do not exceed 2048×2048 pixels, P_{cc} can be any *test image* pixel.

The concatenation of the binary representation of all variables (Table 1) will give a 55 bits long string that represents an individual. Consequently, the search space induced by this representation is $2^{55} \approx 3.6 \times 10^{16}$.

3.3.2. Fitness function

The fitness function is aimed at providing a value (fitness value) to express the individual suitability for the problem

Table 1
Range of encoded variables in an individual

Symbol	Parameter	Range	String length
x	P_{cc} column	[0, (number of image columns—1)]	11 bits
y	P_{cc} line	[0, (number of image lines—1)]	11 bits
sh	Horizontal scale	[0.95, 2.00]	11 bits
sv	Vertical scale	[0.60, 1.50]	11 bits
θ	Rotation angle	[0, 2π]	11 bits

in hand. Since our problem is an image-matching problem, we used F_G (from Eq. (13)) to express the fitness value for some individual I_k :

$$fitness(I_k) = F_G(image_k). \quad (15)$$

However, before using the function $fitness(I_k)$, two critical conditions must be considered. The first condition is illustrated in Fig. 6A, where the enclosing rectangle (defined by I_k) includes the invalid image region (A) and the area (B), which is not an image region. Since the reference matrix values are undefined for the points located in these areas, just the small remaining area in the rectangle inside will influence the $fitness(I_k)$ value. Due to the reduced number of image points to be computed in this case, an unreal fitness value could be assigned to this individual. In this case, $fitness(I_k)$ will be penalized by using the penalty P_1 , defined as follows:

$$P_1(I_k) = n^2 / (n^2 + n_{ip}), \quad (16)$$

where n_{ip} is the number of reference points located in the invalid regions and n^2 is the total number of reference points. When n_{ip} increases, P_1 decreases to lower values in the range [0.0, 1.0].

Another critical condition to be considered is illustrated in Fig. 6B that shows an enclosing rectangle with incomplete length. To be a valid solution, it should include the minimal weld bead length (L) defined by the position marks (0 and 1). Such critical condition happens when the horizontal scale sh is insufficiently large. Therefore, it was decided to penalize $fitness(I_k)$ by using the penalty P_2

$$P_2(I_k) = (sh + 6.sh_{max}) / 7.sh_{max}, \quad (17)$$

where sh_{max} is the maximum value in the sh range (Table 1). Thus, P_2 is maximum (1.00) just when $sh = sh_{max}$.

Using penalty functions P_1 and P_2 , the new fitness function ($fitness'(I_k)$) is defined by the following product:

$$fitness'(I_k) = fitness(I_k) \cdot P_1(I_k) \cdot P_2(I_k). \quad (18)$$

3.3.3. The genetic algorithm working

The initial population of the GA is randomly generated and it is constituted by z individuals (55-bits long strings). Starting from this point, the following procedures describe how a new population can be generated:

- Each decoded individual is evaluated by the fitness function (Eq. (18)).
- The fittest individual is selected, using elitism (as described in [6]), to be inserted in the new population.
- The stochastic tournament selection method (as described in [7]) is used to select pairs of individuals for recombination (crossover).
- With probability p_c , selected pairs of strings are recombined, using uniform crossover operation (as described in [6]), and inserted in the new population.
- Finally, with probability p_m , each individual from the new population have bits (genes) mutated (changed).

Using this procedure, the population is evolved throughout many generations until some stop criterion is met. In this work, we used as stop criterion a maximum number (g) of generations.

Other GA parameters to be set are: population size (z), crossover probability (p_c), and mutation probability (p_m). The tests for parameters setting will be detailed in Section 4.1.

3.4. Object verification

If *object image* i (where $i = 1, 2, \dots, j$) is the result from the genetic search, an hypothesis (**H**) has to be evaluated: *object image* i is an acceptable weld bead sample. This section discusses how such hypothesis can be verified.

Fig. 7 shows samples of non-acceptable (a) and acceptable (b) weld bead detections. The visual criterion of acceptance, used to qualify the detection result, is based in the observation of 4 terms:

- P_{cc} (the central point of reference) must be located in the weld bead bulk.
- Both the weld bead edges cannot be intercepted by the same edge of the enclosing rectangle.
- The enclosing rectangle must include both weld bead edges along its effective stretch (L).
- The height of the enclosing rectangle must be shorter than $3/2$ length of the maximum weld bead thickness.

During the tests for parameters setting (Section 4.1), each time the genetic algorithm was run, hundreds of solutions were randomly generated and evolved through many

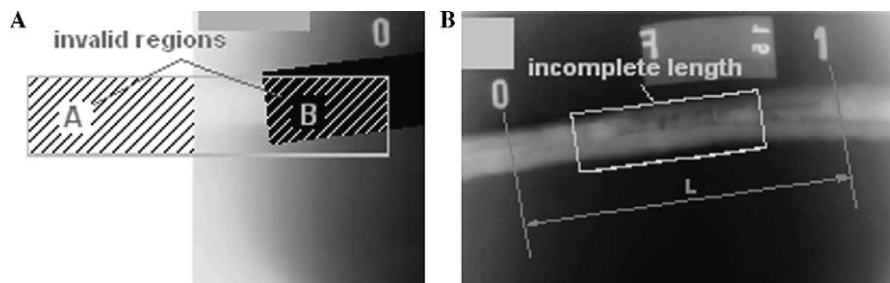


Fig. 6. (A) The enclosing rectangle includes regions (A and B) with invalid reference points. (B) The enclosing rectangle does not include the whole weld length (L) to be inspected.

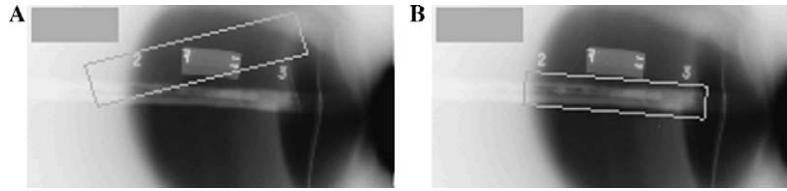


Fig. 7. (A) Bad sample (weld bead detection non-accepted by any term of the visual criterion). (B) Good sample (weld bead detection accepted by all terms of the visual criterion).

generations for each tested image. We only selected the high scored individual from each generation during the genetic evolving process. After discarding duplicated solutions, we classified them into bad or good samples according to the 4 terms we stated as a visual criterion of acceptance. These solutions totaled 97 good and 62 bad samples.

In the chess pieces case [4], the GA output was automatically accepted/rejected based in a minimum threshold for the $fitness(I_i)$ value. In this work, in order to find a feasible threshold for automatic verification, all the matching parameters (F , F_{di} , F_{dj} , F_{ph} , F_{pv}) are investigated using the 97 good and 62 bad samples that were selected. The most important results are shown Fig. 8, where the samples distribution is plotted in terms of F and F_{ph} . Based on these results, a minimum threshold $T_{F_{ph}} = 0.862$ can be observed for the matching parameter F_{ph} . Actually, the sample space is divided into two classes: *weld bead* (with $F_{ph} > 0.862$) and *non-weld bead* (with $F_{ph} \leq 0.862$). Note that, if $T_{F_{ph}} = 0.862$ is used, only one false positive (FP) and one false negative (FN) are observed (see Fig. 8).

As pointed out in the beginning of Section 3, the object verification main function is to evaluate the GA output (*object image i*), by accepting or rejecting it. The current section describes an automatic criterion to perform this

evaluation process. Adopting such criterion, the hypothesis H (*object image i is acceptable as a weld bead sample*) is automatically verified using threshold $T_{F_{ph}} = 0.862$ for the matching parameter F_{ph} . When $F_{ph} \leq 0.862$, it means that H is false and, therefore, *object image i* is rejected. Otherwise, *object image i* is accepted and extracted from the *test image* while the remaining image (*test image – object image i*) is feed-backed to the GA, for a further search, i.e., *object image (i + 1)*.

4. Tests and results

The proposed methodology was implemented in C++ programming language on the Microsoft Windows XP platform. As in our previous work [4], some C++ routines from the GALib package [17] were used for the GA implementation.

For our test series, 111 radiographs of welded joint were acquired as described in Section 2. From one, the *model image* (Fig. 2A) was manually extracted. The remaining 110 radiographs were used for testing. From this group, 22 radiographs were selected for the first test series (parameters setting), described in Section 4.1. Later, 110 pipeline weld joint radiographs (including the 22 radiographs from the last test) were used for the system validation test, described in Section 4.2. Some of the radiographs were also

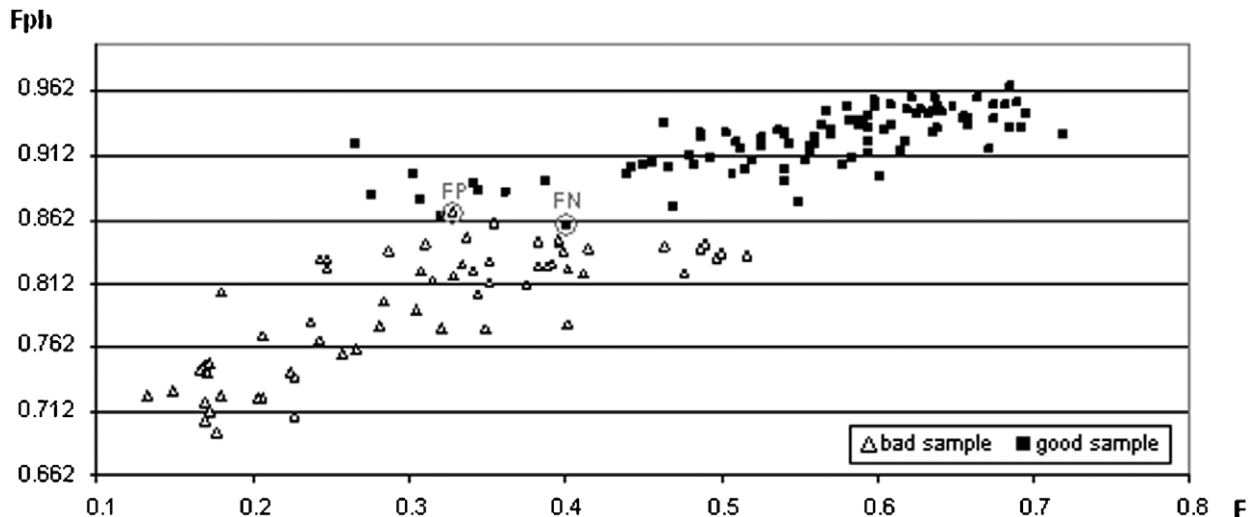


Fig. 8. Distribution of bad and good samples in a plot in terms of the matching parameters F and F_{ph} .

used to generate images with more than one weld bead per image. The test using this last type of images is described in Section 4.3.

4.1. Parameters setting

A number of preliminary tests was done using the group of images composed by the 22 selected radiographs. The objective was to achieve a high level of success in the detection process, without spending a large time in the genetic search. In this way, appropriate (good, but not optimum) values for the GA and the fitness function parameters were defined.

As mentioned in Section 3.3.3, GA parameters to be set are: maximum number of generations (g), population size (z), crossover probability (p_c), and mutation probability (p_m). For the fitness function, the parameters to be set are the binary weights (w_F , w_{Fdi} , w_{Fdj} , w_{Fph} , w_{Fpv}), used in Eq. (13).

4.1.1. GA parameters setting

Initially, weights w_F , w_{Fdi} , w_{Fdj} , w_{Fph} , w_{Fpv} were set to the unity: (1, 1, 1, 1, 1). Using tournament selection method (with tourney size $k = 2$), single-individual elitism, uniform crossover (with $p_c = 0.8$, $p_c = 0.9$, and $p_c = 1.0$), mutation (with $p_m = 0.02$ and $p_m = 0.05$) and $g = 500$, we tested many values in the range [100...750] for the population size (z). The results were analyzed based in the maximum fitness value and the genetic search convergence, by observing the genetic search-evolving plot (see Fig. 9) for each one of the 22 test images.

As a result, convergence to higher fitness values were obtained by using $p_m = 0.02$, $p_c = 0.9$ and at least $z = 500$. Keeping track of the best fitness value throughout generations, a convergent behavior to a high-fitness solution was observed before reaching 100 generations for the 22 images we tested. Such result suggests that a reduced number of generations ($g = 100$ instead of 500) would be enough as stopping criterion. Using these parameter values, the processing time for a run was about 45 s, with a Pentium-IV 2.0 GHz computer.

4.1.2. Testing the relevance of the matching parameters

When all weights (w_F , w_{Fdi} , w_{Fdj} , w_{Fph} , w_{Fpv}) are set to the unity (1, 1, 1, 1, 1), all matching parameters (F , F_{di} , F_{dj} , F_{ph} , F_{pv}) have the same influence in the fitness function value. Thus, in order to check the relevance of each parameter, other sets of weights were tested. Table 2 shows the most important results for this analysis.

According to Table 2, the worst results were obtained when w_F and w_{Fph} were set to zero (test 2 and 5). Such results suggest that the matching parameters F and F_{ph} are very relevant for the weld bead detection problem.

In test 7, when w_F and w_{Fph} were set to 1 and the remaining weights were set to 0, the number of correct detections was the same as before (test 1), when all weights were set to 1. Actually, these results suggest that the parameters F_{di} , F_{dj} , and F_{pv} are not so relevant for the detection problem.

In tests 8 and 9, the influence of the parameters F and F_{ph} were also individually tested. But, the best performance was not kept using just one of them. As a conclusion, it is supposed that both, F and F_{ph} , have almost the same relevance for the weld bead detection problem.

The two unacceptable detections obtained in test 7 are shown in Figs. 10A and B. Both results are unacceptable based in term (c) of the visual criterion of acceptance (Section 3.4). Despite this, such results are very close to what would be expected in case detection was correct. These results encourage us to keep the same parameter values in the validation tests.

4.2. First validation tests

For this test series, 110 pipeline weld joint radiographs (including the 22 radiographs from the last test) were used. As described in Section 2, each radiograph has only one weld bead to be detected. Some results have already been shown in Figs. 5B, 7A and B, and 10A and B. Further results are shown in Table 3.

Using the visual criterion of acceptance (described in Section 3.4), the system response was classified into four classes: false positive (FP), false negative (FN), true

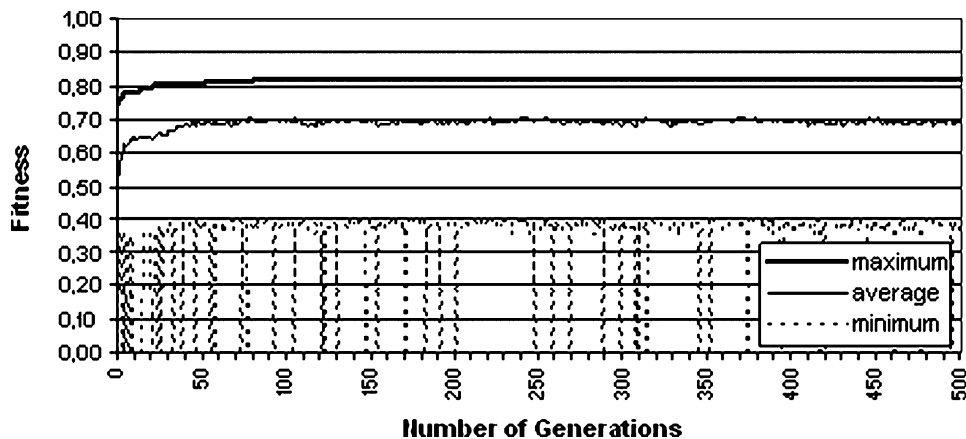


Fig. 9. The maximum, average, and minimum fitness value of population at each generation.

Table 2
Results for different sets of matching parameters

Test	Weight set	Parameters	Non-detections	Unacceptable detections	Correct detections
1	(1, 1, 1, 1, 1)	$F, F_{di}, F_{dj}, F_{ph}$, and F_{pv}	2	0	20
2	(0, 1, 1, 1, 1)	F_{di}, F_{dj}, F_{ph} , and F_{pv}	6	4	12
3	(1, 0, 1, 1, 1)	F, F_{dj}, F_{ph} , and F_{pv}	2	5	15
4	(1, 1, 0, 1, 1)	F, F_{di}, F_{ph} , and F_{pv}	2	3	17
5	(1, 1, 1, 0, 1)	F, F_{di}, F_{dj} , and F_{pv}	4	8	10
6	(1, 1, 1, 1, 0)	F, F_{di}, F_{dj} , and F_{ph}	2	3	17
7	(1, 0, 0, 1, 0)	F , and F_{ph}	0	2	20
8	(1, 0, 0, 0, 0)	F	3	3	16
9	(0, 0, 0, 1, 0)	F_{ph}	2	4	16

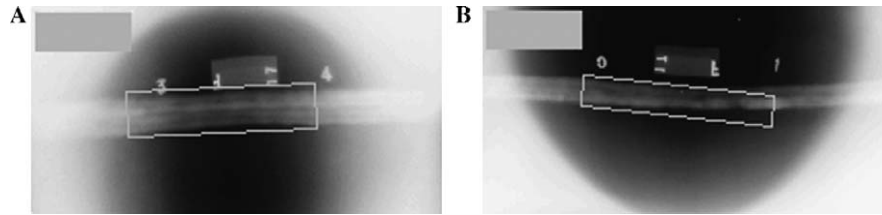


Fig. 10. (A) The length of the detected weld bead is incomplete. (B) The weld bead detection missed some part of the weld bead edge along the effective bead length (L).

positive (TP) or true negative (TN). These classes are described as follows:

- FP: something is detected and recognized as a weld bead, but it is not correct (ex.: Figs. 7A, 10A and B).
- FN: there is a weld bead, but even detected it is not recognized as a weld bead (ex.: Table 3 test image 1).
- TP: there is a weld bead and it is correctly detected and recognized as a weld bead (ex.: Figs. 5B, 7B and Table 3 test image 2).
- TN: there is no weld bead and nothing is recognized as a weld bead.

For test image 1 (Table 3), the weld bead was partially detected, but it was not recognized as a *weld bead* by the verification process. Therefore, it is a case of non-detection, i.e., a false negative (FN). In the case that it was recognized as a *weld bead*, it would be a false positive (FP).

For test image 2 (Table 3), the first detection was correctly recognized as a *weld bead* (TP) and the second detection, that is not a true weld, was correctly rejected (TN). If the second detection was recognized as a *weld bead*, it would be a false positive (FP), since there is just one weld bead in the image.

The accuracy rate (Ar) was evaluated by Eq. (19) using the number of events related to each class (FP , FN , TP , and TN), and shown in Table 4. Using Table 4, one can compute other performance measures, commonly used in data classification literature [8]. These measures, namely sensitivity (Se) and specificity (Sp), mean, respectively, the proportion of real weld bead cases that the system actually classifies correctly, and the proportion of non-weld beads that the system can reject. Sensitivity and specificity are defined in Eqs. (20) and (21), respectively.

$$Ar = (TP + TN) / (FP + FN + TP + TN), \quad (19)$$

$$Se = TP / (TP + FN), \quad (20)$$

$$Sp = TN / (TN + FP). \quad (21)$$

4.3. Second validation test—with different number of radiographs in the same image

In [12], X-ray film strips of about 3.5 in. wide by 17 in. long were digitized four at a time using a scanner. During digitalization, the X-ray filmstrips were positioned side by side in the scanner. Consequently, the welds stay parallel one to another, so that the image lines perpendicularly intercept them. When these procedures are adopted, the image line intensity plot contains all the welds transversal profiles, making possible using the [12] method for the weld detection.

However, in case the welds are not parallel but perpendicularly one to another, or in case the weld bead angle is completely unknown, some welds cannot be correctly detected by that method. To show that our method works even in these cases, some tests were done using images with different number of radiographs that were positioned parallel and perpendicularly in the same image.

Table 5 specifies the number of images, number of weld radiographs per image, FP , FN , TP , TN , and the system performance measures Se , Sp , and Ar . According to the results the accuracy rate was 96% with only one FP case. As shown in Fig. 11, all four welds were detected but one of the results (the unique FP case) is non-acceptable by the visual criterion (Section 3.4).

The results from Table 5 show that the system we proposed is able to effectively detect how many weld beads there are in the image to be analyzed. By comparing the

Table 3
Two test images and the results from genetic search and verification

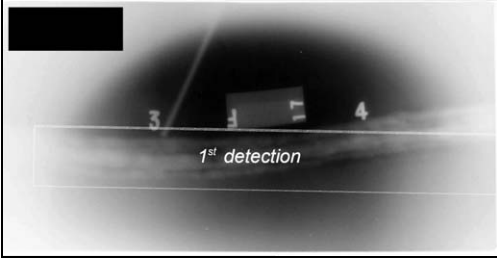
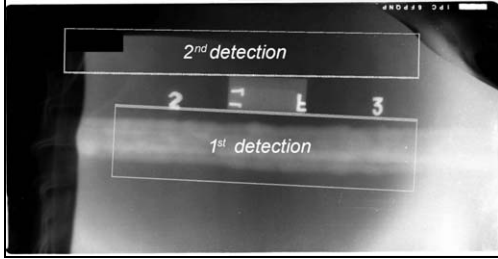



<p>Test Image 1</p> 	<p>Test Image 2</p> 
<p>1st genetic search result (<i>object image 1</i>)</p>  <p>Verification result Not accepted as a <i>weld bead</i>. ($Fph = 0,821669 < T$)</p>	<p>1st genetic search result (<i>object image 1</i>)</p>  <p>Verification result Accepted as a <i>weld bead</i>. ($Fph = 0.947319 > T$)</p>
	<p>2nd genetic search result (<i>object image 2</i>)</p>  <p>Verification result Not accepted as a <i>weld bead</i>. ($Fph = 0.833113 < T$)</p>

Table 4
Confusion table and the values for the performance measures

Actual classes	Test results		Performance measures	
	Weld bead	Non-weld bead		
Weld bead	$TP = 102$	$FN = 2$	$Se = 98.08\%$	$Ar = 94.22\%$
Non-weld bead	$FP = 11$	$TN = 110$	$Sp = 90.91\%$	

results from Tables 5 to 4 it can also be observed that the system accuracy was a little better in the second validation test (96% against 94.22%). However, these results do not necessarily mean that the system performance will always be better for images with multiple weld beads. To check

this performance accurately, a large set of data would be needed.

By observing the weld bead shape from the second (from left to right) weld in Fig. 11 as well as the weld from test image 1 in Table 3, it can be seen that the shape of such weld beads, that were not correctly detected, are more curved than all others. We cannot claim that this was the only reason for non-100% of success in the weld detection process, but such cases suggest that the proposed method is more adequate for detection of weld bead stretches that are linear or almost linear shaped.

To better evaluate the system performance, we regarded the results from both validation tests to construct the new confusion table and calculated the general performance

Table 5
Results obtained in the test with multiples weld beads per image

	Number of images	Number of weld radiographs per image	Results classes				Performance measures		
			FP	FN	TP	TN	Se (%)	Sp (%)	Ar (%)
	2	2	0	0	4	2	100	100	100
	1	3	0	0	3	1	100	100	100
	3	4	1	0	11	3	75	100	93.33
Total	6 images	19 weld radiographs	1	0	18	6	85.71	100	96

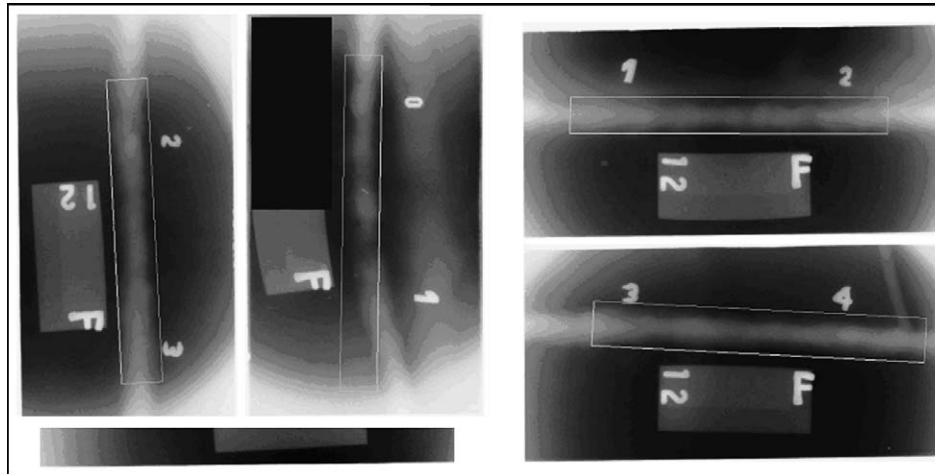


Fig. 11. Image with 4 weld joint radiographs and results that were detected and recognized as weld beads.

Table 6
Confusion table and the values for the performance measures including the both validation tests results

Actual classes	Test results		Performance measures	
	Weld bead	Non-weld bead		
Weld bead	$TP = 120$	$FN = 2$	$Se = 98.36\%$	$Ar = 94.40\%$
Non-weld bead	$FP = 12$	$TN = 116$	$Sp = 90.62\%$	

measures, as shown in Table 6. The accuracy rate (Ar) 94.4% was the general system accuracy.

5. Conclusions

A methodology for detecting, recognizing, and extracting weld beads from digital radiographic images was developed, implemented, and tested. In this methodology, a reference matrix represents the pixel distribution in a single model image with the weld bead in the horizontal position. Five matching parameters (F , F_{dis} , F_{dj} , F_{ph} , F_{pv}), related to this reference matrix, were tested for image matching. As a first conclusion, based in empirical tests, it was observed that parameters F and F_{ph} are very relevant for the problem in hand.

To perform the search for the radiographic image regions most similar to the model (a weld bead), a GA-based procedure was efficiently used. Based on an empirical threshold value for the matching parameter F_{ph} , GA results are classified as *weld bead* or *non-weld bead*. For each image, the number of solutions generated and evaluated by the GA-based system was 5×10^5 (population size \times number of generations) and that was done in only 45 s (using a Pentium-IV 2.0 GHz). Comparing the number of evaluations with the total search space size (2^{55}), the GA-based method can find a satisfactory solution by testing less than $10^{-9}\%$ of the search space, with an efficiency (general accurate rate) of 94.40%, as shown by the results (Table 6).

Our obtained results strongly suggest that this method is feasible for detecting position, width, length, and angle of

weld beads in digital radiographic images. By testing images with different number of radiographs per image, we also showed that the proposed system is able to detect and count how many welds are there in the image, even if one weld bead is perpendicular to another and the number of welds is previously unknown. However, this method has a limitation: it was observed that the proposed method is more adequate for detection of weld bead stretches that are linear or almost linear shaped instead of curved welds.

This work is part of a project in which the automatic inspection of weld joint radiographs is the final goal. Currently, the presented weld recognition system has already been used for the automatic linear weld bead extraction. Future work will focus on curved welds, defect detection and classification of the defect types.

Acknowledgments

This work has been supported by CNPq and (ANP) Petroleum National Agency—ANP/MCT (PRH10-CE-FET-PR). The software developed in this work used the GALib genetic algorithm package, written by Matthew Wall at the Massachusetts Institute of Technology.

References

- [1] American Society of Mechanical Engineers-ASME Boiler and Pressure Vessel Code, Section V, Subsection A, Article 2: Radiographic Examination, 2001.
- [2] B. Bhanu, J. Peng, Adaptive integrated image segmentation and object recognition, IEEE Trans. Syst. Man Cybernet. 30 (4) (2000) 427–441.

- [3] L. Bocchi, L. Ballerini, S. Hässler, A new evolutionary algorithm for image segmentation, in: *Lecture Notes in Computer Science*, vol. 3449, 2005, pp. 264–273.
- [4] T.M. Centeno, H.S. Lopes, M.K. Felisberto, L.V.R. Arruda, in: *Object detection for computer vision using a robust genetic algorithm*, in: *Lecture Notes in Computer Science*, vol. 3449, 2005, pp. 284–293.
- [5] O. Cordon, S. Damas, E. Bardinet, 2D Image registration with iterated local search, in: *Advances in Soft Computing—Engineering, Design and Manufacturing*, Proc. of NSC7, 2003, pp. 1–10.
- [6] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA, 1989.
- [7] D.E. Goldberg, B. Korb, K. Deb, Messy genetic algorithms: motivation, analysis, and first results, *Complex Syst.* 3 (5) (1989) 493–530.
- [8] D.J. Hand, *Construction and Assessment of Classification Rules*, John Wiley, New York, 1997.
- [9] S.E. Lawson, G.A. Parker, Intelligent segmentation of industrial radiographic images using neural networks, in: *Proc. SPIE Architectures and Systems Integration III*, vol. 2347, 1994, pp. 245–255.
- [10] T.W. Liao, J. Ni, An automated radiographic NDT system for weld inspection: part I—weld extraction, *NDT&E Internat.* 29 (3) (1996) 157–162.
- [11] T.W. Liao, K. Tang, Automated extraction of welds from digitized radiographic images based on MLP neural networks, *Appl. Artif. Intell.* 11 (13) (1997) 197–218.
- [12] T.W. Liao, D. Li, Y. Li, Extraction of welds from radiographic images using fuzzy classifiers, *Informat. Sci.* 126 (2000) 21–40.
- [13] T.W. Liao, Classification of welding flaw types with fuzzy expert systems, *Expert Syst. With Appl.* 25 (2003) 101–111.
- [14] H.I. Shafeek, E.S. Gadelmawla, A.A. Abdel-Shafy, I.M. Elewa, Assessment of welding defects for gas pipeline radiographs, *NDT&E Internat.* 37 (4) (2004) 291–299.
- [15] R.R. Silva, L.P. Calôba, M.H.S. Siqueira, J.M.A. Rebello, Pattern recognition of weld defects detected by radiographic test, *NDT&E Internat.* 37 (6) (2004) 461–470.
- [16] K.S. Simunic, S. Loncaric, A genetic search-based partial image matching, in: *Proc. ICIPS'98, IEEE Internat. Conf. on Intelligent Processing Systems*, 1998, pp. 119–122.
- [17] Wall, M., 2003. GAlib A C++ library of genetic algorithm components vs. 2.4.5. <<http://lancet.mit.edu/ga/>>.