Automatic segmentation of radiographic images in industrial applications

PIOTR BANIUKIEWICZ

*Department of Electrical and Computer Engineering,
West Pomeranian University of Technology,
Szczecin, Poland*

e-mail: baniuk@zut.edu.pl

Abstract: A technology that utilizes penetrating rays is one of the oldest nondestructive testing methods. Nowadays, the process of radiogram analysis is performed by qualified human operators and automatic systems are still under development. In this work we present advanced algorithms for automatic segmentation of radiographic images of welded joints. The goal of segmentation of a radiogram is to change and simplify representation of the image into a form that is more meaningful and easier to analyse automatically. The radiogram is divided into parts containing the weld line, image quality indicators, lead characters, and possible defects. Then, each part is analysed separately by specialized algorithms within the framework of the Intelligent System for Radiogram Analysis.

Key words: radiography; nondestructive testing; image processing

1. Introduction

Defects and discontinuities in material can reduce the strength of the element and can lead to catastrophic failure. This is critical especially in the case of welding due to the application of this technique for joining stress-exposed parts such as ship hulls, skyscrapers, bridges, and others. Many different energy sources can be used for melting materials. The most popular are gas flame, electric arc, laser beam, friction, and ultrasound [1]. Welding can be done in many different environments, including open air, underwater, and space. The chemical and physical processes that occur during melting and joining of materials are very complicated and not fully understood. For example, metal cooling is a highly nonlinear physical process which proceeds simultaneously with chemical reactions. These processes result in increased mechanical tensions in material as well as forming new chemical compounds in the joint area (e.g. gases, slag, etc.). These undesired products are often the source of potential cracks and discontinuities in the joint. Additionally, the heat applied to the joint area in the process of welding causes overheating of the surrounding material, resulting in changes in its mechanical

properties. This is a potential source of unwanted signals for all nondestructive testing (NDT) methods. Thus, there is a need to develop techniques that make the crack detection process more reliable and accurate.

A technology that utilizes penetrating rays is one of the oldest NDT methods. Over the years this technique has been the most important NDT method. Fast and accurate x-ray converters that have been developed recently open up new possibilities for x-ray applications and make this technique more flexible than ever before. Digital radiography combines modern digital image processing algorithms with the traditional x-ray testing method. Fast development of digital radiography has resulted in the emergence of new NDT technology.

2. Problem statement

One of the biggest problems in assuring the demanded quality of a product is the need to analyse a large number of radiograms. In the shipbuilding industry this number can reach thousands. All radiograms must be analysed by a proficient radiologist in order to find those welds which do not meet very restrictive safety norms. This process takes a lot of time and effort but it tends to be fully automated. Advanced computerized systems for radiogram analysis are still under development [2] and decisions taken by human beings are still considered to be the most reliable and trustworthy. But the importance of systems that support or even superseding of radiologist, such as automatic radiogram analysis systems or automatic defect recognition systems, is growing. The real radiograms taken on a production line are far from ideal. Of course they must meet very restrictive and precise norms concerning the quality of the image, whether digital or traditional, but the location and orientation of objects in the picture can vary. For example in practice it is hard to keep the same position of welds in all radiograms or alignment of markers such as lead letters and digits (practically no one pays attention to this aspect).
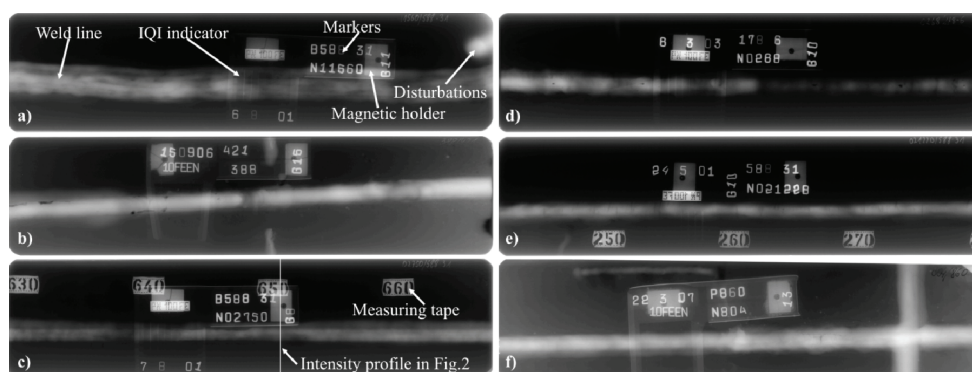

Fig. 1. Examples of real radiograms of welds

Figure 1 shows examples of real radiograms of welds from the shipbuilding industry obtained from Technic-Control company [3]. They all meet technological norms and can be used

to determine the quality of the weld but it is clearly visible that they differ in position of weld and other objects as well as in brightness and contrast. The human operator is able to analyse and describe the radiogram by finding some characteristic points, objects, and indicators commonly contained in it. For example, the usual method of measuring image quality in radiography is the use of Image Quality Indicators (IQIs). The process of analysis is completely intuitive for an experienced operator but for a computer it can be a very demanding task lying in the fields of image analysis, object recognition, and computer vision. The main problems are:

- Varying appearances of radiograms even when acquired during the same process (Fig. 1).
- Presence of various objects, which can disturb detection of other objects (Fig. 2a).
- Very low contrast of some objects (e.g. wire indicators) (Fig. 2b).
- Welds intersections (Fig. 1f).

Figure 2c shows that additional "technological" objects (typically made from lead) usually put on the investigated specimen are the source of signals having higher intensity than the signal under investigation. In this way they disturb the process of automatic analysis.
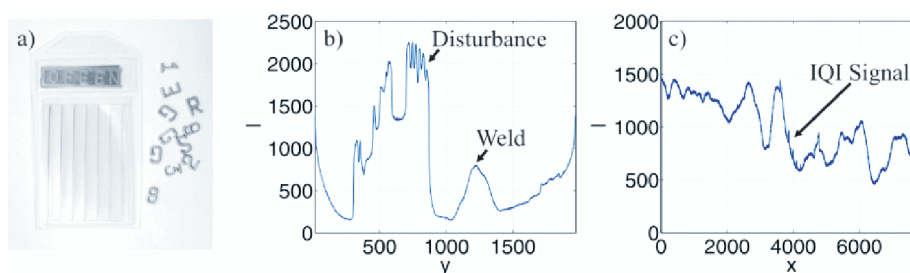
Fig. 2. (a) Lead letters and IQI wire indicator; (b) the intensity profile taken from the radiogram presented in Figure 1c along the y axis; (c) the intensity profile taken from Figure 1a along the weld line

The common method of overcoming these difficulties is to apply specialized algorithms targeted at performing only one task, for example finding the IQI on the weld or detecting flaws in the weld. This is the main reason for developing methods of radiogram segmentation. In this paper the author presents algorithms for extracting selected objects from real radiograms. The objects under investigation are: the weld line, description characters, and IQIs. After segmentation, extracted objects are passed to other algorithms for further analysis. The weld line is analysed by defect detection algorithms and the IQI is used for estimation of the image quality value. Characters and letters are recognized to give an enhanced radiogram description.

3. Description of the segmentation methods

The segmentation is performed in three steps. In the first step the weld line is detected and extracted. In the second step the position of the IQI is obtained. Finally, all remaining objects

are localized and recognized. Assuming that only the *V* butt joint geometry is taken into consideration, the cross section of the weld has a Gaussian shape (Fig. 2c) with small linear trend. This assumption is the basis of the weld line detection algorithm. The algorithm utilizes robust approximation to track the weld in the radiogram. The intensity function *I(y)* of each column *x* of the radiogram, which is a weld profile, is approximated separately [4] by the sum of the Gaussian function and linear function (1),

$$I(y)_x = a \cdot e^{\left(\frac{-(y-b)}{c}\right)^2} + d \cdot y + f \, , \tag{1}$$

where *y* is the row number. Up and down bounds of parameters *a*, *b*, *c*, *d* and *f* are set to prevent unexpected changes in function parameters in approximated columns. The function parameters of the previous weld profile are used as the start parameters of the next one. Moreover, a vector of weights **w** associated with the response data *I(y)* has been introduced. The weight *w(y)* defines the contribution of point *y* to fitted curve *I(y)*. The vector $\mathbf{w_x}$ for column *x* is estimated to be an average of approximations of *N* previous columns, scaled to the range 0-1. The initial values of coefficients *a*, *b*, *c*, *d* and *f* for each approximation *I(y)* are medians of *N* previous values of these coefficients. The experimental results show that the median operator works much better, making the process of approximation less susceptible to local signals of high intensity such as characters and other welds. Finally, the maxima of Gaussian functions *I(y)$_x$* define the centre of the weld, whereas the edges of the weld are assumed to be 5% of the height of *I(y)$_x$*. In the last step the centre of the weld and its edges are smoothed by means of a low-pass filter to eliminate small high-frequency variations. The initial values of **w** used for approximation of the first column are obtained from *N* columns of the raw radiogram whereas *a*, *b*, and *c* are estimated during the approximation process without any limitations. Examples of results of weld detection are presented in Figure 3.
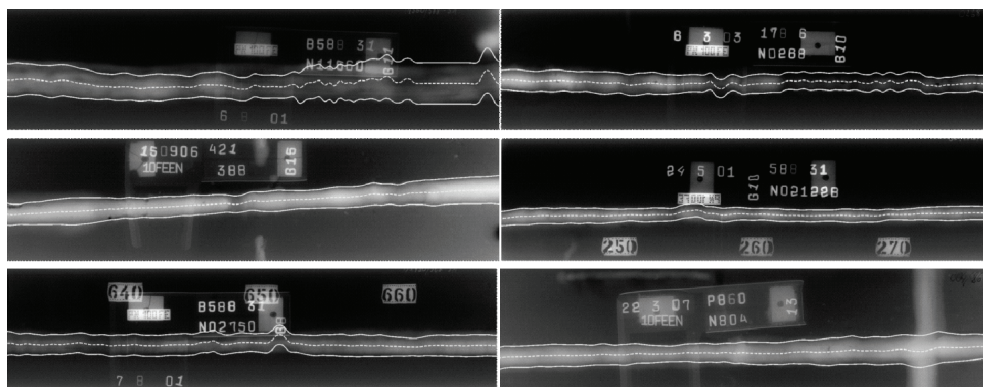

Fig. 3. Result of automatic weld detection for radiograms from Figure 1

This method works correctly even in the presence of high-intensity objects near the weld line. Resistance to possible interferences mainly depends on the number of preceding steps

that are used to estimate current parameters. In the case of a clean radiogram the weld can be successfully approximated with $N = 50$ whereas for more disturbed pictures the parameter N should be greater than 200. The critical aspect for efficiency of the algorithm is selection of initial N columns which forms the base of the tracking algorithm. In order to achieve good approximations $I(y)_x$ of the weld, the initial columns used as initial weights in (1) should contain only a clear weld profile without any other signals. After analysis of a database of radiograms containing over 90 pictures, it has been observed that in most cases good representatives can be obtained from the beginning or end of the radiogram because the main region of interest (commonly containing all unwanted objects) is located in the centre of the picture.

A typical wire IQI contains six or seven straight wires of the same material as the specimen or similar material and having various diameters. The wires are held parallel to one another in a plastic mount with appropriate identification symbols. The most important problem relating to IQI detection is the low signal amplitude of the IQIs. Moreover, a typical radiogram suffers from a nonlinear trend that results from nonhomogeneity of the specimen and its non-uniform thickness. The amplitude of the trend signal is much greater than the intensities of particular wires (Fig. 2c). In the approach proposed here, only the presence and position of IQIs is evaluated. Under the fundamental initial assumptions for the algorithm lies the norm [5] that regulates the placement of IQIs in the radiogram. In the case of welds, it is expected that IQIs are always located on the weld line and almost perpendicular to it. Thus, the region of the analysis is limited to the weld line and a small surrounding area. The plastic case covering the wires in the IQI is the best visible part of the whole indicator in the radiogram as well as in the intensity profile. This fact is used in the process of detection and localization of IQIs.

The algorithm works with averaged intensity profiles I taken along the weld line. The position of an IQI is obtained from a fusion signal F arising from two methods commonly used in the pattern recognition. These methods are: 1) measuring the dependence between the pattern sought p and the signal I using the Pearson product-moment correlation coefficient (2) and 2) evaluating the phase correlation between the Fourier transform of pattern p and the Fourier transform of signal I. The position of the maximum of F corresponds with the position of IQI in the radiogram. The correlation coefficients are calculated between the pattern p and the window of size n sliding over the signal I with a step of 1,

$$r_i = \frac{n\sum I_i p - \sum I_i \sum I p}{\sqrt{n\sum I_i^2 - \left(\sum I_i\right)^2}\sqrt{n\sum p^2 - \left(\sum p\right)^2}}, \qquad (2)$$

where I_i stands for the i-th position of the window in the signal I and n is the number of elements of p and I_i.

The pattern p is generated artificially as an average calculated from dozens of intensity profiles of IQIs obtained from various radiograms. The resulting waveform can be assumed to be the best representative of an IQI signal holding general features of that signal. Figure 4 shows examples of results of the IQI detection obtained for the radiogram depicted in Figure 1a.
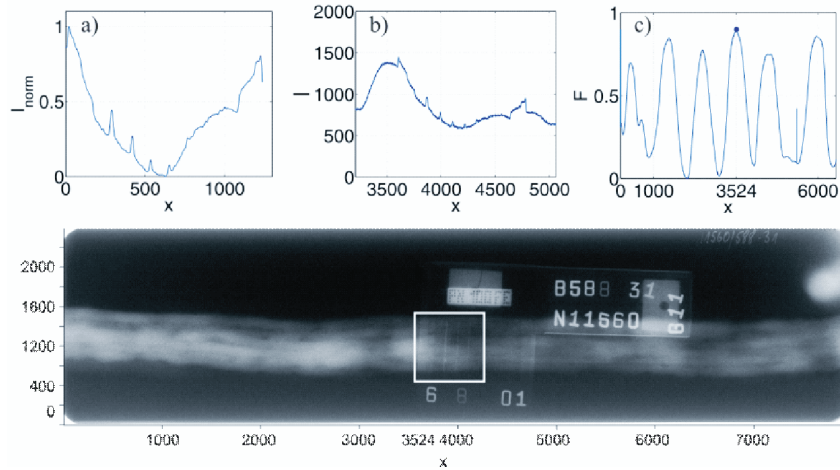
Fig. 4. a) Plot of normalized pattern *p* representing a general IQI signal; b) fragment of an intensity profile along the weld line from Fig. 1a containing an IQI; c) plot of fusion function *F*. Below: the radiogram with the detected IQI

Finally, the letter recognition algorithm is applied to the radiogram in order to read all the characters and technological descriptions placed on it. Localization of the weld line obtained in the previous step allows recognition to be sped up by excluding it from the analysis. Moreover, after exclusion of the weld line, lead characters become the most intense objects in the picture, which facilitates the process of thresholding and selecting individual letters. In this process the image is binarized using the Sauvola thresholding method [6, 7], and then selected morphological operations such as removing isolated pixels, opening, erosion, and indexation are applied in order to make objects separable. This process is depicted in Figure 5.

Fig. 5. Results of successive stages of letters extraction: a) fragment of radiographic image; b) image after thresholding using Sauvola algorithm; c) image after filtering; d) image after morphological closing; e) image after morphological eroding; f) image after indexation

The recognition process depends on the comparison of selected characters with known objects included in a database. The database is a vast collection of various characters, letters, and symbols obtained from real radiograms delivered by a company operating in the field of NDT. It contains over 3000 characters in 18 groups relating to the following symbols: the digits 0-9 and the letters F, E, N, G, P, R, D, and B. The comparison of the selected object with those from the collection is difficult and its results are ambiguous mostly because the

same character can be found in the image at different scales and rotations. To overcome this problem all selected objects as well as those in the database are processed by the Fourier-Mellin (FM) transform [8], which makes them invariant to translation, scale, and rotation. The FM transform is a useful tool commonly applied for image recognition. The Fourier Transform itself is translation invariant and its conversion to log-polar coordinates converts the scale and rotation differences to vertical and horizontal offsets that can be measured. A second FFT, called the Mellin transform, gives a transform-space image that is invariant to translation, rotation, and scale. Thus, all objects are transformed to FM amplitude spectrums of size 256×256. The essential features of transformed images of the same characters are similar and independent of their rotation and scale (Fig. 6).
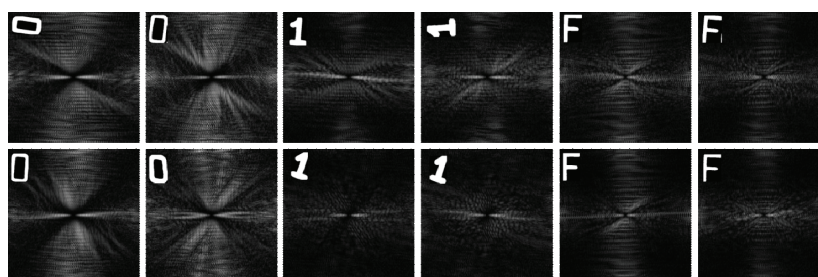

Fig. 6. Selected representatives of groups of characters and theirs FM transforms independent of scale and rotation

By treating a picture as a vector in 256×256 dimensional space it is possible to classify all pictures from the database according to how they are grouped in the input space. Due to differences between characters it is expected that vectors representing them will be well-separated in multi-dimensional space. The application of the FM transform ensures that objects representing the same character but differing in scale or rotation will be assigned to the same class. The self-organizing feature map (SOFM) has been chosen as the classifier. The applied neural network had 65536 inputs and a hexagonal topology based on a 6-by-6 neurons grid. This yields 36 clusters. After the training process, in the ideal case every cluster should contain only characters belonging to the same group of symbols. However, in a real case every cluster is quantitatively dominated by the same symbols but also contains a small number of other symbols whose presence represents the recognition errors. The reliability *RE* achieved by the classifier, defined as the percentage of properly recognized characters for every group, is as follows:

Table 1. Results of recognition of lead symbols in the radiograms

Group	RE [%]	Group	RE [%]	Group	RE [%]	Group	RE [%]
0	91.28	5	58.65	F	43.25	R	47.23
1	96.25	6	12.23	E	62.22	D	2.36
2	72.64	7	25.53	N	98.89	B	15.62
3	86.81	8	77.56	G	82.26		
4	86.56	9	4.56	P	32.44		

Unfortunately, there are pairs of characters that are similar to each other such as G and 6, B and 8, and D and 0. This meaningfully reduces the effectiveness of the classifier and is the source of recognition errors.

4. Conclusions

The four methods presented allow selected objects typically found in radiograms to be located and recognized. Thus, further analysis of these objects can be performed by targeted algorithms working on previously selected regions of interest. This approach minimizes disturbances and improves the performance and operational reliability of the targeted methods. The weld detection algorithm shows accuracy in locating straight welds as well as slightly skewed or bent ones. The robustness of the algorithm mainly depends on the ratio between the number of preceding lines *N* and the length of disturbances as well as on the distance between the weld and other objects. The proposed algorithms are embedded in the intelligent system for radiograms analysis which the authors are involved in creating.

References

[1] Cary, Howard B., *Modern Welding Technology*, Prentice Hall, ISBN: 0130309133 (2001).
[2] European Commission sponsored project *Development of novel digital radiography technology. To facilitate the traditionally less research intensive inspection industry sector change from manual film radiography to automated digital.* Contract No NMP2-C-2005-515746.
[3] http://www.technic-control.com.pl/
[4] Sikora R., Baniukiewicz P., Chady T., et al. *Comparison of selected weld defect extraction methods.* Review of Progress in Quantitative Nondestructive Evaluation, AIP Conference Proceedings 975: 1034-1041 (2008).
[5] European norm EN 462, part 1. *Non-destructive testing – Image quality of radiographs – Part 1: Image quality indicators (wire type) – Determination of image quality value.* (1994).
[6] Mehmet Sezgin, Bülent Sankuj, *Survey over image thresholding techniques and quantitative performance evaluation*, Journal of Electronic Imaging 13(1): 146-165 (2004).
[7] Jaakko Sauvola, Matti Pietikäinen, *Adaptive document image binarization.* Pattern Recognition 33: 225-236 (2000).
[8] Derrode S., Ghorbel F., *Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description.* Computer Vision and Image Understanding 83(1): 57-78 (2001).