AUTOMATIC LABEL REMOVAL FROM DIGITIZED WELD RADIOGRAPHS

SOO SAY LEONG[1], MANI MARAN RATNAM[2]*, ZAHURIN SAMAD[3] &
MOHD ASHHAR KHALID[4]

Abstract.   This paper presents a methodology to remove labels automatically from digitized weld radiographs as part of the automatic weld defect detection process. An algorithm was developed to detect and remove labels printed onto weld radiographs before weld extraction algorithm or defect detection algorithm is applied. Normality test was used to determine if the intensity profile parallel to the weld contains label pixels. Thresholding followed by region filling operations were carried out to remove the labels. The algorithm was tested on 50 weld radiographs with labels and the labels on 90% of these images were successfully removed.

Keywords: Weld radiography, label removal, normality test

Abstrak.   Kertas kerja ini membentangkan keadah mengeluarkan label secara automatik daripada radiograf kimpalan sebagai sebahagian daripada proses pengesanan kecacatan automatik. Suatu algoritma telah dibangunkan untuk mengesan dan mengeluarkan label yang tercetak pada radiograf kimpalan sebelum algoritma penyarian kimpalan dikenakan. Ujian kenormalan digunakan untuk menentukan sama ada profil keamatan selari dengan kimpalan mengandungi piksel label. Pengambangan diikuti dengan operasi mengisi kawasan dilakukan untuk mengeluarkan label. Algoritma tersebut diuji ke atas 50 radiograf kimpalan yang mempunyai label dan label pada 90% daripada imej tersebut dikeluarkan dengan jayanya.

Kata kunci: Radiograpfi kimpalan, pengeluaran label, ujian kenormalan

1.0   INTRODUCTION

A radiograph is a photographic record produced by the passage of *X*-rays or Gamma rays through an object onto a film. The image formed on the film is digitized and converted into two-dimensional array of picture elements (pixels) before computer-based enhancement or automatic analysis is carried out.

Radiographs are usually marked to clearly indicate the weld number or other traceable identification [1]. Examples of identification that may appear in a radiograph

---

[1,2,3] School of Mechanical Engineering, Universiti Sains Malaysia, Engineering Campus, Sri Ampangan, 14300 Nibong Tebal, Penang, Malaysia.
*Corresponding author: Tel.: +6-04-5937788. Fax: +6-04-5941025.
Email: mmaran@eng.usm.my

[4]   Intelligent Systems Group, Malaysia Institute of Nuclear Technology Research (MINT), 43650 Bangi, Selangor, Malaysia.
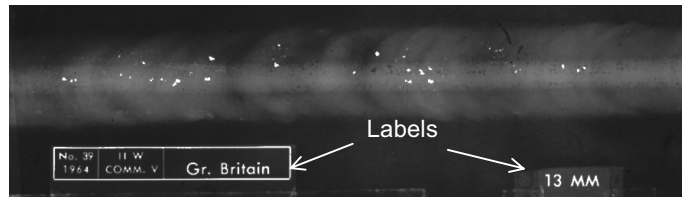
Figure 1   Labels on a weld radiograph

are film number, part number, name and/or logo of the film manufacturer, date of examination etc. [2]. In addition, image quality indicators (IQI) are used on the film to provide information regarding contrast sensitivity and definition of the radiograph. These identifications and IQIs appear in the film as labels as shown in Figure 1.

Although the labels provide important information concerning a radiograph, it is desirable to remove them before an automatic weld defect detection and interpretation algorithm is applied to the weld region. This will ensure that subsequent weld extraction and defect detection operations will not be affected by the pixels belonging to the labels. Although much effort has been made on weld extraction [3-5] and defect detection [6-9], no previous attempt has been made to remove the labels from weld radiographs automatically. In most of the past research, the label regions were removed manually or the defect regions cropped out to carry out the defect identification, e.g. [6, 9-10]. The use of manual intervention to define the regions to be processed or omitted in the weld radiographs is an impediment to the automatic interpretation of the radiographs. In this work, a methodology to remove labels automatically from digitized weld radiographs is proposed.

2.0   METHODOLOGY

The automatic label removal methodology developed in this research is summarized in the flow-chart in Figure 2. Each step of the flow-chart is discussed in detail in the following sub-sections.

2.1   Contrast Enhancement

In this work, reference radiographs from the International Institute of Welding (IIW) were scanned using an *X*-ray film scanner (*Cobrascan CX-612-T*) in 12-bit resolution and saved in *TIFF* image format without compression. Since some of the original radiographs were poor in contrast, image contrast enhancement was carried out before further processing. Linear contrast stretching and histogram equalization are two widely used methods for image contrast enhancement. However, histogram equalization has the undesired effect of over-emphasizing noises [11]. Contrast stretching on the other hand, does not cause this problem. Contrast stretching is a contrast enhancement
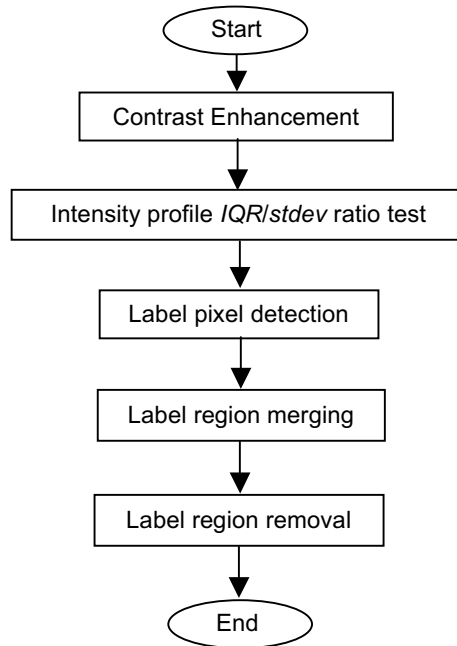
```
                    ┌─────────┐
                    │  Start  │
                    └────┬────┘
                         ▼
          ┌──────────────────────────────┐
          │     Contrast Enhancement     │
          └──────────────┬───────────────┘
                         ▼
      ┌──────────────────────────────────────┐
      │ Intensity profile IQR/stdev ratio test │
      └──────────────────┬───────────────────┘
                         ▼
          ┌──────────────────────────────┐
          │     Label pixel detection    │
          └──────────────┬───────────────┘
                         ▼
          ┌──────────────────────────────┐
          │     Label region merging     │
          └──────────────┬───────────────┘
                         ▼
          ┌──────────────────────────────┐
          │     Label region removal     │
          └──────────────┬───────────────┘
                         ▼
                    ┌─────────┐
                    │   End   │
                    └─────────┘
```

Figure 2   The proposed label removal methodology

technique for linearly mapping the intensity values of an image to a new range. In this work, the intensity value was stretched to a full range of intensity level in an 8-bit image, which is 0 to 255 gray levels.

2.2 Intensity Profile $IQR/stdev$ Ratio Test

The location of label pixels in the image must first be identified before these pixels can be removed. Since the label pixels are located close to the edge of the weld radiograph and well away from the weld region [2], they can be removed from the image without affecting the weld. In the automatic radiograph interpretation it is important to remove the labels so that the weld bead can be extracted automatically using appropriate algorithms. An analysis was carried out on the intensity profiles taken in the i-direction of the weld radiograph (Figure 3) to detect the labels in the image. These intensity profiles were divided into four different types:

Type 1: Intensity profile from a background region without labels.

Type 2: Intensity profile from background region with labels.

Type 3: Intensity profile from weld region without defects.

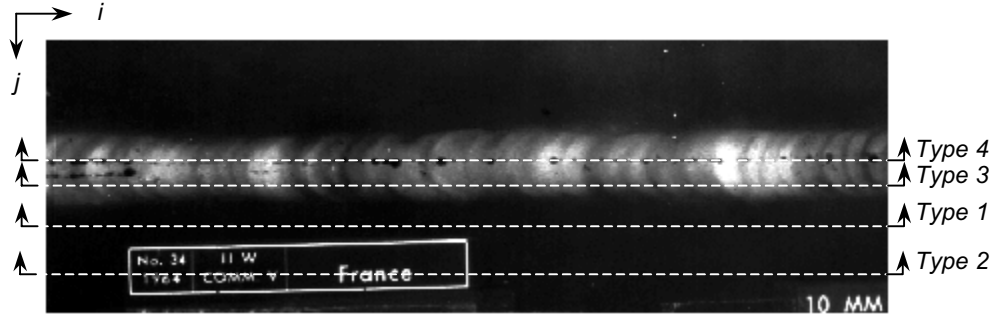*Type 4*: Intensity profile from weld region with defects.

Figure 3   Location of four types of intensity profiles along *i*-direction

The aim is to detect *Type 2* intensity profiles that occur across labels in the image. Normality test was performed on the intensity values of the profile in order to distinguish *Type 2* intensity profile from the others. Several methods can be used for the normality test. One of the tests is performed by calculating the *IQR/stdev* ratio of the pixel intensity data, where *IQR* is the inter-quartile range of the data while *stdev* is the standard deviation of the data. The inter-quartile range, *IQR* is the distance between the upper and lower quartiles given by,

$$IQR = Q_U - Q_L \tag{1}$$

where the upper quartile Q_U for a data set is the 75th percentile while the lower quartile Q_L is the 25 percentile of the intensity profile. If the distribution of the data is normal, the *IQR/stdev* will be equal to 1.34 [12]. Four examples of *IQR/stdev* ratio calculations are illustrated for the four types of intensity profiles in Table 1.

Ten sample profiles of each type, i.e. a total of forty samples, were collected randomly from ten weld radiographs. *IQR/stdev* ratios were calculated for each profile and the results are presented in Table 2. The average value of *IQR/stdev* ratio for *Type 2* profile was found to be lower compared to *Type 1*, *Type 3* and *Type 4* profiles. The difference between the average values of *IQR/stdev* for each type of intensity profile and the value of *IQR/stdev* for a normal distribution profile are 12.58%, 73.78%, 14.68% and 4.81% respectively for *Type 1*, *Type 2*, *Type 3* and *Type 4* profiles. *Type 1*, *Type 3* and *Type 4* profiles were found to have *IQR/stdev* values close (within 20%) to the *IQR/stdev* value for the normal distribution profile. Thus, the presence of label pixels can be identified based upon the *IQR/stdev* ratio due to the difference between the ratio for label pixels and non-label regions.

Based on the results in Table 1, a threshold value for *IQR/stdev* ratio was selected in order to differentiate *Type 2* intensity profile from the other three profiles. Since the average *IQR/stdev* value of ten samples for *Type 1*, *Type 3* and *Type 4* intensity profiles is 1.483 while the average value for *Type 2* of intensity profiles is 0.351, a value between these two ratios, i.e. 0.9, was selected as the threshold. Thus, if the *IQR/stdev* ratio in a

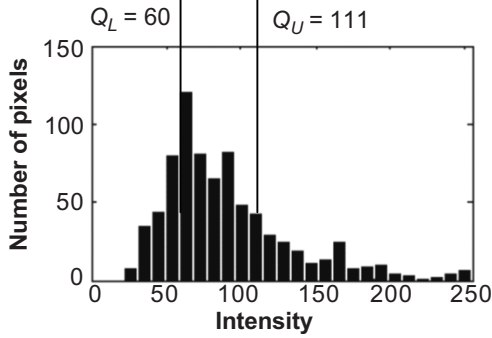**Table 1** Intensity histogram and calculation of $IQR/stdev$ ratio
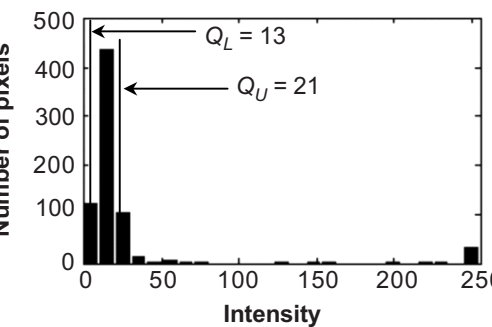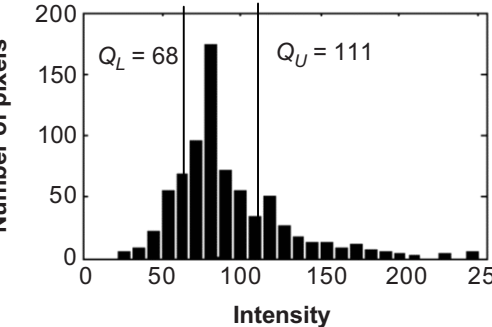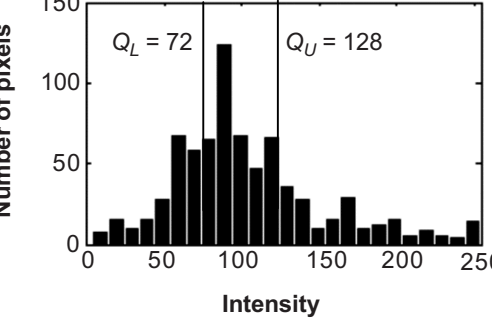
Intensity Histogram	IQR/*stdev* Ratio
	Type 1 Profile $Q_L = 60$ $Q_U = 111$ $IQR = Q_U - Q_L$ $IQR = 51$ $stdev = 44.19$ $\therefore \dfrac{IQR}{stdev} = 1.15$
	Type 2 Profile $Q_L = 13$ $Q_U = 21$ $IQR = Q_U - Q_L$ $IQR = 8$ $stdev = 57.15$ $\therefore \dfrac{IQR}{stdev} = 0.14$
	Type 3 Profile $Q_L = 68$ $Q_U = 111$ $IQR = Q_U - Q_L$ $IQR = 43$ $stdev = 39.11$ $\therefore \dfrac{IQR}{stdev} = 1.10$
	Type 4 Profile $Q_L = 72$ $Q_U = 128$ $IQR = Q_U - Q_L$ $IQR = 56$ $stdev = 50.29$ $\therefore \dfrac{IQR}{stdev} = 1.11$

Table 2   *IQR/stdev* ratios for *Type 1* to *Type 4* intensity profile

Image	*IQR/stdev* ratios			
	Type 1	**Type 2**	**Type 3**	**Type 4**
1	1.582	0.533	1.643	1.527
2	1.002	0.393	1.593	1.076
3	2.040	0.092	1.210	1.749
4	1.558	0.254	1.562	1.312
5	1.126	0.531	1.632	1.903
6	1.398	0.329	1.769	1.150
7	1.570	0.287	1.849	1.047
8	1.822	0.482	1.585	1.555
9	1.114	0.161	1.401	1.492
10	1.875	0.453	1.124	1.234
Average	1.509	0.351	1.537	1.405
Difference[1]	12.58%	73.78%	14.68%	4.81%

[1]Compared to normal distribution profile

new intensity profile is more than 0.9, this implies that there are no label pixels in the profile. On the other hand, label pixels are present in the profile if the *IQR/stdev* ratio is equal to or less than 0.9.

2.3   Label Pixel Detection

From the analysis of *IQR/stdev* ratio on the intensity profiles, we conclude that if the *IQR/stdev* ratio is more than 0.9, all the pixels along the profile are non-label pixels. In this case, these pixels were assigned to 0 gray value. On the other hand, for profiles having *IQR/stdev* ratio less than or equal to 0.9, histogram thresholding operation was used to extract label pixels from the intensity profile. Thresholding is to define a range of brightness values in an image, select the pixels within this range as belonging to the foreground, and reject all of the other pixels to the background [13]. This is illustrated as follows.

Figure 4 shows an example of a weld radiograph with labels. An intensity profile taken across the label along $j = 223$ has a *IQR/stdev* ratio of 0.0949. Since label pixels are detected in this profile (*IQR/stdev* < 0.9), histogram thresholding was carried out to extract these pixels.

The intensity histogram of the profile along $j = 223$ is shown in Figure 5. By using *graythresh* function in Matlab, a threshold value of 55 was obtained. Thus, pixels with intensity value higher than 55 was set to 1 while the remaining pixels were assigned to 0. After the thresholding operation was carried out for all the intensity profiles with *IQR/stdev* ratio less than or equal to 0.9, a binary image was obtained. Figure 6 shows the resulting image after label pixels detection.
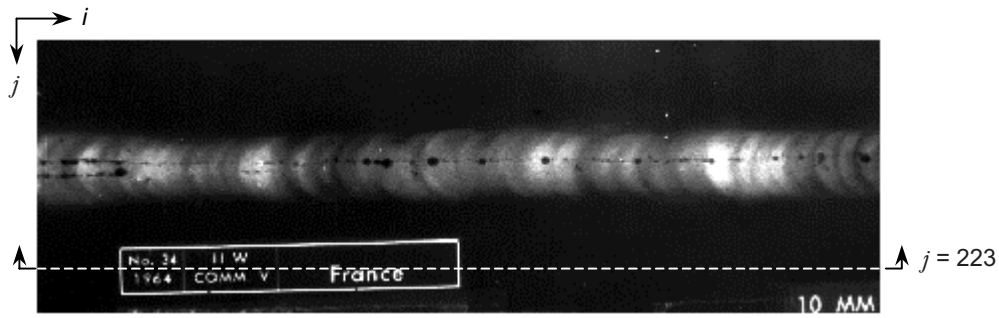
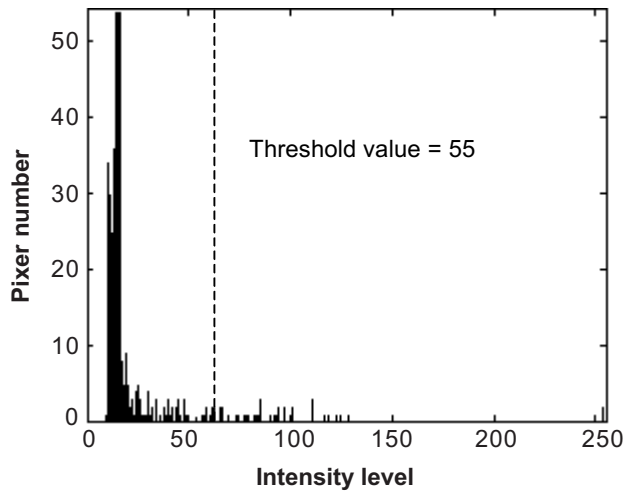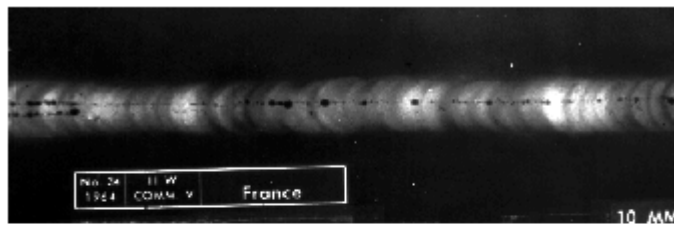**Figure 4**   Intensity profile taken across labels on $j = 223$

Figure 5   Histogram of intensity profile on $j = 223$

(a)

(b)

Figure 6   Example of label pixel detection: (a) is the original image, (b) detected label pixels

2.4   Label Region Merging

The label pixel detection algorithm was able to extract most but not all the label pixels in the image. Therefore, label region merging was carried out to merge all label pixels in the image into connected regions. The detected label pixels that belonged to the same label in the image were merged into one label region. Dilation technique was applied to connect the label pixels. The dilation process involves the use of a convolution mask known as a structuring element, which acts over both external and internal object boundary, depending on the size and shape of the structuring element [14].

A rectangular structuring element was chosen in this research based on the observation that the labels in the weld radiographs are approximately rectangular in shape. In addition, the size of the structuring element also plays an important role. The selection of the size of the structuring element is crucial because an undersized element will not be able to connect all the label pixels, while an oversized element tends to remove useful information from the image. Different sizes of structuring element was tested on the image and the result showed that a 9×19 rectangular window structuring element produced the best result in connecting all the label pixels.

Therefore, dilation operation with 9×19 structuring element was carried out on the detected label pixel binary image shown in Figure 6(b). An image with all merged label regions was obtained as shown in Figure 7, in which the detected label pixels were merged into several separate label regions. In the subsequent stages, these label regions were removed from the image to complete the label removal process.

Figure 7   Merging of labels region for image shown in Figure 6(b)

2.5   Label Region Removal

After merging the label regions, the main label areas in weld radiograph were identified. Feret boxes were constructed enclosing the label regions as shown in Figure 8(a). A feret box is the smallest rectangle orientated according to a specific reference that encloses the shape [15]. The intensity of the pixels located inside the feret boxes was filled using the *roifill* function in Matlab. The *roifill* function smoothly interpolates inward from the pixel values on the boundary of the feret box by solving a Laplace
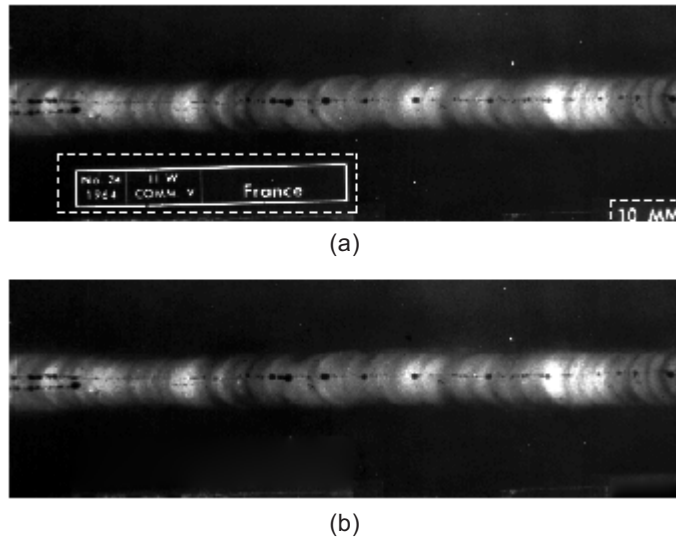
(a)

(b)

Figure 8   Results of label region removal: (a) constructed feret boxes,
(b) removal of labels pixels using *roifill*

equation [14]. Figure 8(b) shows the result after removing the label region and filling
the region using the *roifill* function.

3.0   RESULTS AND DISCUSSIONS

The proposed label removal algorithm was tested on 50 weld radiographs digitized
using an *X*-ray film scanner. These radiographs comprised 45 having several types of
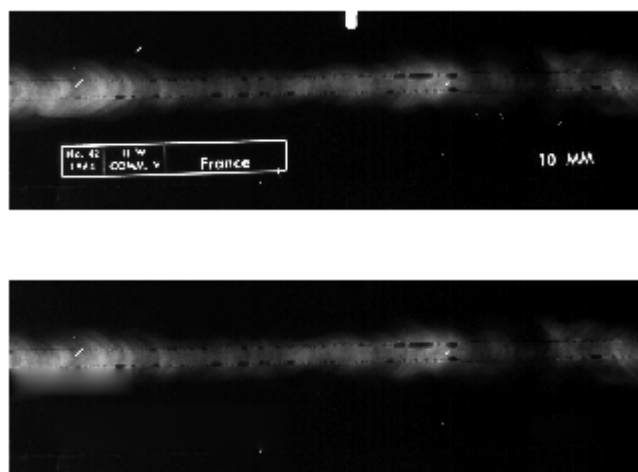weld defects and 5 defect-free images. Out of 50 digitized radiographs, labels on 45

Figure 9   Results of label removal algorithm (Sample 1): (a) original image
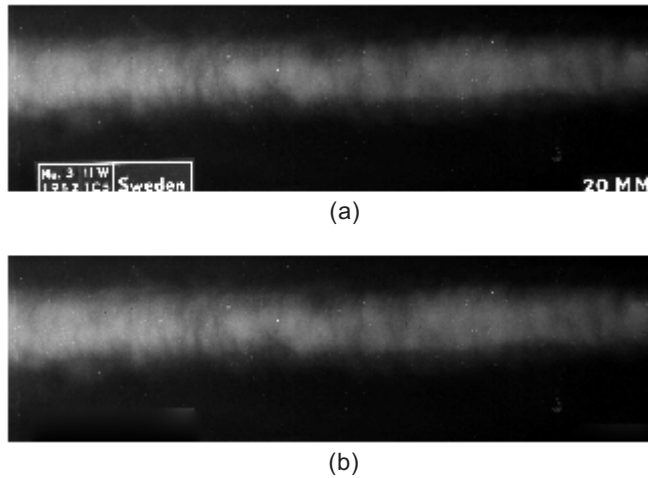and (b) image after label removal

(a)

(b)

Figure 10   Results of label removal algorithm (Sample 2) (a) original image
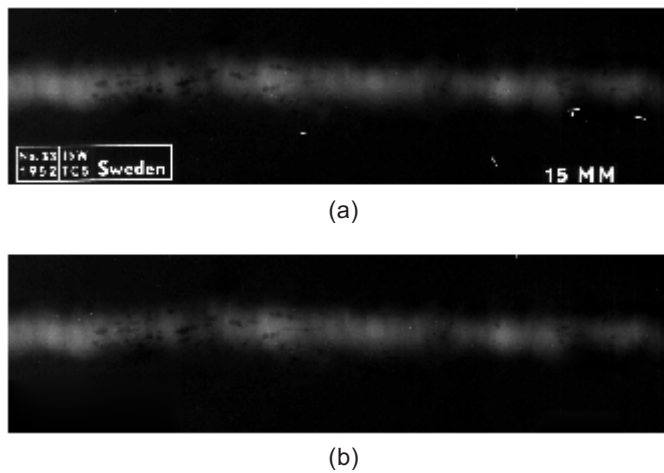and (b) image after label removal

(a)

(b)

Figure 11   Results of label removal algorithm (Sample 3) (a) original image and
(b) image after label removal

were successfully removed. Figures 9 to 12 show some examples of weld radiographs
where labels on the image were successfully removed.

However, out of the 50 weld radiographs tested labels on 5 images were not
completely removed, such as the one shown in Figure 13. For this example, the
algorithm was only able to remove labels located at the lower side of the weld region
but not those located at the upper side of the weld.

The labels located at the upper side of the weld in Figure 13 were not removed
because the *IQR/stdev* ratio of the pixels on the intensity profiles taken across these
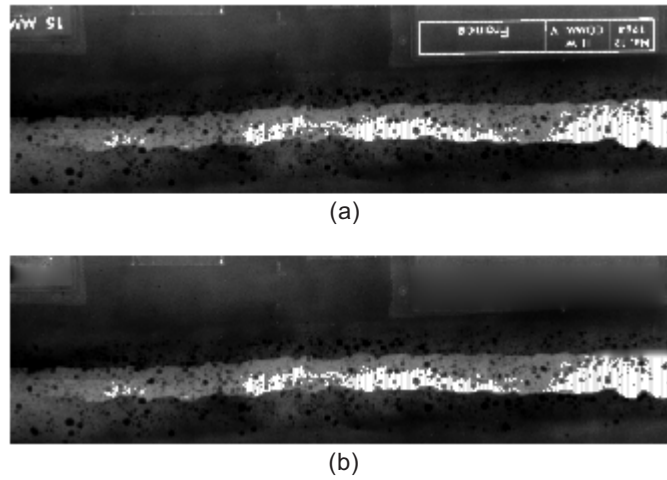labels is higher than 0.9. For instance, the intensity profile taken across the labels

(a)

(b)

Figure 12   Results of label removal algorithm (Sample 4) (a) original image
and (b) image after label removal
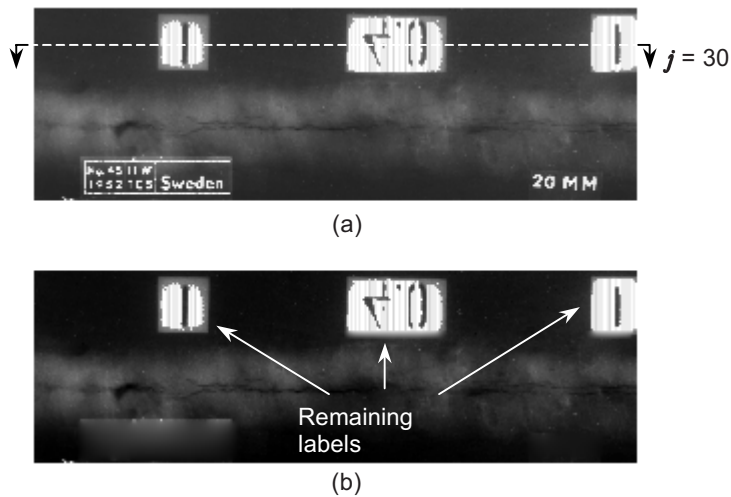

(a)

(b)

Figure 13   Example of labels remaining after application of label removal algorithm:
(a) radiograph with labels and (b) labels remaining after the application of
label removal algorithm

along $j = 30$ has an *IQR/stdev* ratio of 1.13. Since this ratio is higher that the threshold (0.9), the intensity profile was misinterpreted as a non-label profile. When the intensity profile is taken across large labels, a large quantity of label pixels is present in the profile. In this case, the *IQR* value increases due to the large intensity range between label pixels and background pixels. The increase of *IQR* value leads to an increase of *IQR/stdev* ratio. Consequently, a high *IQR/stdev* ratio was noticed for intensity profiles taken across large labels, which leads to inaccuracy in the label detection algorithm.

Therefore, the proposed technique can only remove labels that have *IQR/stdev* ratio less than or equal to 0.9. Since the labels in 90% of the images were detected and removed successfully due to their low *IQR/stdev* ratios, the proposed algorithm can be implemented in automatic weld radiograph interpretation systems for removing labels.

4.0 CONCLUSION

A label removal algorithm was developed to remove labels from digitized weld radiographs in the early stage of the automatic weld defect detection process. The algorithm applies the normality test (*IQR/stdev* ratio) on pixel intensity profiles taken parallel to the weld in the digitized image to detect label pixels. The detected label pixels that belong to the same labels on the image were merged by dilation operation to form label regions. Finally, the labels were removed using the *roifill* function in Matlab.

The algorithm was tested on 50 weld radiographic images. Out of the 50 images, labels on 45 images were successfully removed. However, labels on 5 of the images were not completely removed. Large labels on these five images produced *IQR/stdev* ratios above the selected threshold during the normality test. Since intensity profiles taken across the labels in most radiographs have low *IQR/stdev* ratios, this technique can be effectively used as a pre-processing stage in the automatic analysis of weld radiographs. Once the labels in a radiograph are removed, weld extraction algorithms can be applied to the whole radiograph without the need to segment the weld region manually.

ACKNOWLEDGEMENT

REFERENCES

[1]   American Bureau of Shipping. 2002. *Guide for Nondestructive Inspection of Hull Welds.* American Bureau of Shipping. Houston. USA.

[2]   Chuse, R. and B. E. Carson. 1993. *Pressure Vessels - The ASME Code Simplified.* McGraw-Hill Inc. USA.

[3]   Liao, T. W. and J. Ni. 1996. An Automated Radiographic NDT System for Weld Inspection: Part I- Weld Extraction. *NDT&E Int.* 29(3): 157-162.

[4]   Lawson S. W. and G. A. Parker. 1994. Intelligent Segmentation of Industrial Radiographs Using Neural Networks. Proceeding of SPIE 2347: 224-255.

[5]   Liao, T. W., D. M. Li, and Y. M. Li. 2000. Extraction of Welds From Radiographic Images Using Fuzzy Classifiers. *Information Sciences* 126:21-40.

[6]   Liao, T. W. and Y. M. Li. 1998. An Automated Radiographic NDT System for Weld Inspection: Part II - Flaw detection. *NDT & E Int.* 31(3): 183-192.

[7]   Just, T., W. Thale, and R. Clausen. 1998. Interpretation of Radiographs by Digital Image Processing. 7th European Conference on Non-Destructive Testing 3(1) (*http://www.ndt.net/article/ecndt98/nuclear/205/205.htm*).

[8]   Bonser, G. and S. W. Lawson. 1998. Defect Detection in Partially Completed SAW and TIG Welds Using Online Radioscopy and Image Processing. Proceedings of SPIE: Process Control and Sensors for Manufacturing. 3399: 231-239.

[9]   Wang, G. and T. W. Liao. 2002. Automatic Identification of Different Types of Welding Defects in Radiographic Images. *NDT&E Int*. 35: 519-528.

[10]   Shafeek, H. I., E. S. Gadelmawla, A. A. Abdel-Shafy, and I. M. Elewa. 2004. Assessment of Welding Defects for Gas Pipeline Radiographs Using Computer Vision. *NDT&E Int*. 37: 291-299.

[11]   Chang, D. C. and W. R. Wu. 1998. Image Contrast Enhancement Based on a Histogram Transformation of Local Standard Deviation. *IEEE Transactions on Medical Imaging*. 17(4): 518-531.

[12]   Mendenhall, W. and T. Sincich. 1995. *Statistics for Engineering and the Sciences*. Upper Saddle River, New Jersey: Prentice Hall.

[13]   Russ, J. C. 1999. *The Image Processing Handbook*. 3rd edition. USA: CRC Press,.

[14]   MathWorks. 2001. *Image Processing Toolbox: User's Guide*. Verison 3. The MathWorks, Inc.

[15]   Costa, L. F. and R. M. Cesar. 2001. *Shape Analysis and Classification*. CRC Press. USA.