

Classification of weld defects based on the analytical hierarchy process and Dempster–Shafer evidence theory

Hongquan Jiang¹ · Rongxi Wang^{1,2} · Zhiyong Gao¹ · Jianmin Gao¹ · Hongye Wang¹

Received: 8 June 2016 / Accepted: 16 October 2017 / Published online: 26 October 2017
© Springer Science+Business Media, LLC 2017

Abstract Weld quality management is currently one of the most concerning issues in the manufacturing industry. In this paper, a novel method is proposed for weld defect classification based on the analytical hierarchy process (AHP) and Dempster–Shafer (DS) evidence theory. First, to overcome the problem of traditional DS methods, which weigh every feature equally in classification, a method is proposed based on AHP to calculate the weight of features (WF) of a weld defect, which can then be utilized in classification. Then, an improved method based on DS evidence theory is presented to improve the accuracy of classification, which includes calculation of the standard value of features based on frequency histograms analysis and an improved Dempster’s rule for combination based on WF. A case study on the classification of steam turbine weld defects is provided to illustrate and evaluate the proposed techniques. The results show that the proposed method increases the correct recognition rate of classification with limited samples, making DS evidence theory applicable to weld defect classification.

Keywords Radiographic testing · Classification · Analytical hierarchy process (AHP) · Dempster–Shafer (DS) Evidence theory

Introduction

In the manufacturing industry, radiographic testing (RT) is an important non-destructive testing technique for finding defects present inside a material. Radiographic images used in the classification of weld defects are of great significance for weld quality management, reliability, and large-scale equipment safety (Zahran and Al-Nuaimy 2002; Silva and Mery 2007a). Additionally, the results of weld inspection provide useful information for identifying potential problems in the fabrication process, which is necessary for improving welding operations (Liao 2009; Zapata et al. 2012). Therefore, improving the efficiency and accuracy of weld defect classification is an important technical problem.

In current weld quality management, many efforts have been devoted for developing an automated inspection system to improve the accuracy and efficiency of RT (Silva and Mery 2007b; Liao 2003). An automated RT system usually includes the following four steps (Liao 2009; Shen et al. 2010), as shown in Fig. 1. (1) Image preprocessing mainly includes noise reduction and image enhancement. Noise reduction smoothes the RT images, and image enhancement improves the contrast and highlights the defects and details in RT images. Image enhancement usually includes contrast enhancement and edge enhancement methods (Gu et al. 2013; Mu et al. 2013) such as the histogram transformation method (Zahran et al. 2013), gray level transformation method (Gao and Hu 2014), and fuzzy enhancement method (Movafeghi 2015). (2) Defect segmentation finds and isolates weld defect regions from the rest of the radiographic

✉ Rongxi Wang
rongxiwang@163.com

Hongquan Jiang
jhqxjtu@163.com

Zhiyong Gao
zhgyao@mail.xjtu.edu.cn

Jianmin Gao
gjm@mail.xjtu.edu.cn

Hongye Wang
wanghongye@stu.xjtu.edu.cn

¹ State Key Laboratory for Manufacturing System Engineering, Xi’an Jiaotong University, Xi’an, China

² Institute of Manufacturing Systems and Quality Engineering in the Department of Mechanical Engineering, Xi’an Jiaotong University, Xi’an, China

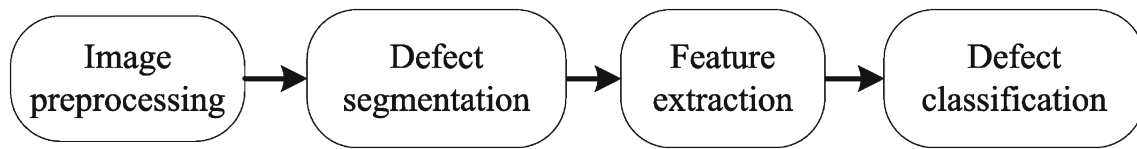


Fig. 1 Process of automated RT system

image, providing the foundation of feature extraction and defect classification in automated RT systems (Nacereddine et al. 2006, 2013). (3) Feature extraction obtains a set of features—generally including edge-based and region-based features—that can describe the characteristics of the weld defects (Shen and Gao 2010). (4) Defect classification, i.e., defect pattern classification, which classifies defects according to the extracted defect features, is the focus of this paper.

It should be noted, there are many other methods relevant to weld quality management, such as thermal image analysis, ultrasonic testing (UT), and pulsed thermography (PT). For example, Sreedhar et al. (2012) proposed an online weld-monitoring system based on thermal image analysis via infrared cameras for tungsten inert gas (TIG) welding. Salchak et al. (2016) proposed a method for classifying possible defects with corresponding dimension limits based on UT. Du et al. (2008) developed a UT-based weld defect classification method by integrating a discriminant tracing algorithm and a probabilistic neural network. Zhu et al. (2017) developed an improved feature extraction algorithm for eddy current PT to realize automatic defect identification. RT, UT, and PT each have limitations, e.g., defects lying in certain planes cannot be detected by some of the methods. Because RT is widely used in current practical applications, especially in the turbine manufacturing industry, this paper focuses on RT and its application.

In recent decades, many studies have been performed on the classification of weld defects (Silva and Mery 2007a, b), and several methods have been applied to the classification of defects: fuzzy systems, artificial neural networks (ANNs), and support vector machines (SVMs) (Zhang et al. 2005). For example, Liao (2003) proposed a fuzzy expert system approach for defect classification, for which the most important step is acquisition of fuzzy rules. Zapata et al. (2010) proposed an adaptive-network—based fuzzy inference system for classification of weld defects. Fuzzy systems have a tradeoff between accuracy and interpretability, i.e., the accuracy can be improved but at the cost of interpretability. Wang and Liao (2002) proposed a method based on a multi-layer perceptron (MLP)-ANN with 12 features for defect classification. Lim et al. (2007) selected an optimal set of features from 25 features using a statistical approach and then used an MLP-ANN to classify the defects. ANNs can achieve good performance with a large number of training samples, but they can make judgment errors with only a small number

of training samples. Zhang et al. (2005) utilized asymmetrical SVMs to classify weld defects. Li (2009) employed an SVM to classify a local defect embedded in a homogeneous copper-clad laminate surface. A direct multiclass SVM (DMSVM) for classifying defects has also been proposed (Shen et al. 2010). The SVM-based methods have good generalization, but their performance depends on the selected defect samples and defect features. In addition, there are some methods that integrate different classification theories. For example, You et al. (2015) developed a laser weld defect classification method by integrating a feed-forward neural network and an SVM. Fan et al. (2016) proposed a solder bump recognition method integrating fuzzy theory and an SVM. Mu et al. (2013) proposed a weld defect classification method based on principal component analysis (PCA) and an SVM, which is more effective than traditional SVM methods. Generally, these methods belong to machine learning/data mining techniques, and classification accuracy is greater in such techniques. Ultimately, all of the aforementioned methods cannot give the probability that a weld defect would belong to a certain type owing to the small differences between some types of weld defects.

Dempster–Shafer (DS) evidence theory (Pohl and Genderen 1998), which can address the uncertainty associated with classification, is currently being introduced to the field of pattern recognition. For example, Hong et al. (2012) proposed a method for fault diagnosis based on a modular neural network and DS theory. Liu et al. (2011) introduced an approach for glass defect identification based on DS evidence theory. Gao et al. (2012) applied DS evidence theory to fuse different features in image subcategory classification. DS evidence theory has also been applied to fault condition diagnosis (Pan et al. 2011), which can combine multiple features to decide the fault condition type. In fact, the classification of weld defects is a process of multi-information fusion, i.e., DS theory can be applied to fuse the different weld defect features (henceforth called evidence) for classification.

However, there are several problems in applying DS to weld defect classification. The first problem is the calculation of weighting factors for each feature when using features as the evidence, i.e., each feature of a defect has a different contribution for classification, whereas traditional methods assign features equal weighting. To handle this problem, the analytic hierarchy process (AHP) is used in this work. AHP

(Saaty 1980, 1990) is a decision method that can split the elements related to the decision into different hierarchies such as goals, principles, and programs (Maruthur et al. 2013). Furthermore, AHP combines qualitative analysis with quantitative analysis, so it can effectively solve complex decision problems whose objects and criteria are difficult to quantify (Hafizan et al. 2016). In order to promote the scientific nature of weight prediction and avoid the contradiction between weight prediction and actuality due to subjectivity, AHP can be used to calculate the weight of features (WF) before applying a DS-based method. The second problem is the standard value of feature (SVF) identification for elements in the frame of discernment. The value of each feature (e.g., length, aspect ratio) for a certain type of defect varies, so it should be decided before using the DS method. The third problem is the calculation of the combined rules for multi-evidence cases. This issue is crucial to the success of applying DS methods; in contrast, traditional methods do not utilize weighting factors for features, so they cannot be directly applied to the classification of weld defects.

To overcome the aforementioned problems, a novel method based on AHP and DS evidence theory is proposed for weld defect classification. First, a method based on AHP is proposed to calculate the WF for weld defects. Then, an improved method based on DS evidence theory that includes calculation of the SVF based on frequency histogram analysis (FHA) and an improved Dempster’s rule of combination based on the WF is presented. Finally, a case study involving the classification of steam turbine weld defects is provided to illustrate the proposed techniques. This method enables application of DS theory to weld defect classification, and it classifies weld defect types with better performance than traditional methods.

The paper is organized as follows. “Feature selection” section describes the features that are used in this paper. “Calculation of WF based on AHP” section discusses the WF based on AHP. “Improved DS method for weld defect classification” section introduces the improved DS method for weld defect classification. “Weld defect classification case study” section describes the weld defect classification case study. “Results” section presents the results of the case study. “Discussion” section discusses how the proposed method compares to other classification methods. Finally, “Conclusion” section presents the conclusions on the proposed method and gives suggestions for future research.

Feature selection

As highlighted in Fig. 2, there are generally five types of weld defects: porosity (PO), slag inclusion (SL), lack of penetration (LP), lack of fusion (LF), and crack (CR). The features of weld defects can be categorized into two classes: edge-based

features and region-based features. Edge-based features are principally centered on measured geometric properties, and region-based features are focused on the intensity characteristics. As shown in the white boxes in Fig. 2, the weld defect area should be segmented before feature extraction and feature selection. In this paper, seven defect features were selected from previous works (Shen et al. 2010; Mu et al. 2013), as shown in Table 1. These features can be used for defect classification based on DS theory, i.e., the features can be described as evidence and fused to determine the class of a weld defect.

Calculation of WF based on AHP

In traditional classification methods, features are usually weighted equally, and the obtained feature values are directly used for defect classification. However, in the process of defect classification, the role of each feature varies with the defect type, i.e., different features have different weights in defect classification. If the same weight is used for each feature, the classification accuracy may be reduced. In contrast, assigning different weights to different features according to the defect type improves the classification accuracy. Therefore, the WF should be determined before applying the DS-based method. In this paper, a method based on AHP for calculating the WF is proposed. The specific process of this method is as follows:

- (1) The evaluation index w_i ($i = 1, \dots, 7$), which represents the weight value of weld defect features, is determined. For example, w_1, w_2, \dots, w_7 can be defined such that they represent the weight values of L, Ar, \dots, K_u , respectively, as shown in Table 1.
- (2) The comparison matrix A is constructed as follows:

$$A = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \dots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \dots & \frac{w_2}{w_n} \\ \dots & \dots & \dots & \dots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \dots & \frac{w_n}{w_n} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \tag{1}$$

where $a_{ij} = w_i/w_j$ represents the pair-wise comparison of index w_i relative to index w_j . For simplicity, the value of a_{ij} was defined directly in this study, rather than from w_i , and a_{ij} was set based on a 1-to-9 scale of relative importance, which is also called the Saaty scale. Odd numbers (1, 3, 5, 7, 9) represent equal importance, moderate importance, strong importance, very strong importance, and extreme importance, respectively, and even numbers (2, 4, 6, 8) represent intermediate values between adjacent judgments (Hafizan et al. 2016).

Fig. 2 Defect samples: **a** porosity (PO), **b** slag inclusion (SL), **c** lack of penetration (LP), **d** lack of fusion (LF), and **e** crack (CR)

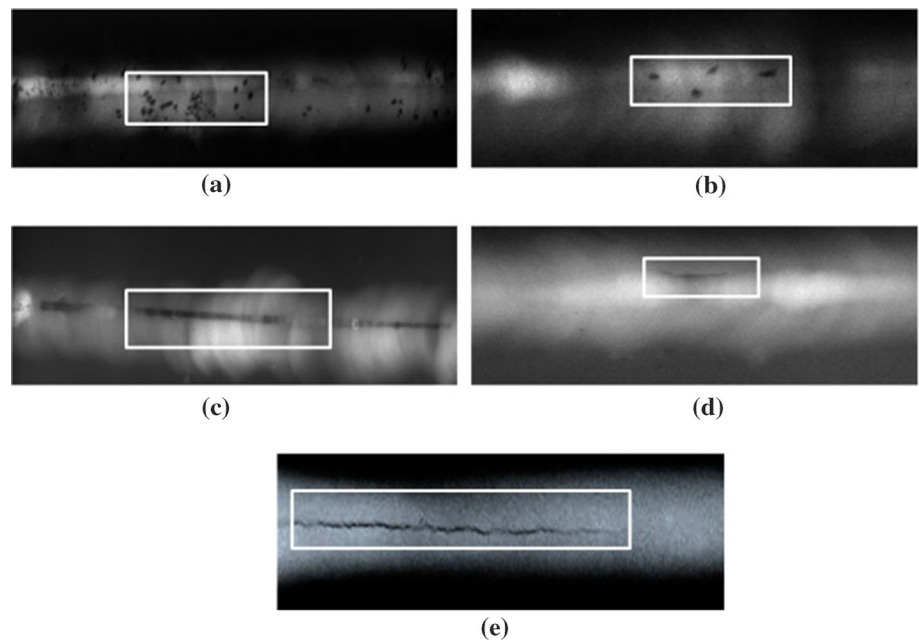


Table 1 Defect features

	Feature	Feature type	Formula
L	Length	Edge	–
A_r	Aspect ratio	Edge	$A_r = \frac{L}{H}$
S_p	Sharpness	Edge	$S_p = \frac{A_1 + A_2}{A}$
R_e	Roughness of defect edge	Edge	$R_e = \frac{1}{L} \sum_{i=1}^L h_{i+1} + h_{i-1} - 2h_i $
R_r	Roughness of defect region	Region	$R_r = \frac{1}{A} \sum_{i=1}^{L-1} \sum_{j=1}^{h_i-1} \nabla^2 f $
S_k	Skewness	Region	$S_k = \frac{1}{L} \sum_{i=0}^L \sum_{k=0}^{h_i} \left \frac{2(x_3^k + x_1^k - 2x_2^k)}{(x_3^k - x_1^k)} \right $
K_u	Kurtosis	Edge	$K_u = \frac{1}{L} \sum_{i=0}^L \sum_{k=0}^{h_i} \frac{x_3^k - x_2^k}{x_4^k - x_1^k}$

Other features are described in References (Shen et al. 2010; Mu et al. 2013)
 H width of defect area, L length of the defect area, A area of the defect ($= L \times W$)

(3) In order to test the consistency of markings in matrix A , the random consistency ratio (CR) is calculated as acceptable if $CR < 0.1$. Otherwise, the comparison matrix should be revised.

$$CR = \frac{CI}{RI} \tag{2}$$

(4) If $CR < 0.1$, according to Eqs. (2) and (3), the matrix weight values, w_i , are

$$w_i = \frac{v_i}{\sum_{i=1}^n v_i} \tag{4}$$

In the formula, the coincident indicator (CI) is

$$CI = \frac{\lambda_{\max} - n}{n - 1} \tag{3}$$

where $v_i = \sqrt[n]{\prod_j a_{ij}}$. Therefore, the list of weights is

$$W_{\text{index}} = \{w_1, w_2, \dots, w_n\} \tag{5}$$

where λ_{\max} is the maximum eigenvalue of A and RI is a reference value, i.e., an indicator of the impact of category i on the reference area. The CR ranges from zero to one, and it can be used to determine the consistency of the judgment entries. It is widely agreed that the comparison matrix, A , is

Thus, W_{index} can be used as the WF during weld defect classification.

Selecting features and determining relationships between features are key steps when using AHP. If the selected features are not reasonable, or the relationships between features are not correct, as shown in Eq. (1), the results of AHP will be inaccurate. Thus, to ensure the rationality of the hierarchy and the accuracy of defect classification, the selected features and the relationship between features should be reasonable and correct.

Improved DS method for weld defect classification

Brief introduction of DS evidence theory

DS theory is an important method in uncertainty reasoning first introduced by Dempster (1968) and later extended by Shafer (1976); it plays vital roles in the representation and fusion of uncertain information and the study of random sets (Han et al. 2014). DS theory is widely used in fields such as information fusion, pattern recognition, and decision analysis. DS theory is defined as follows.

If Θ represents a finite set of elements, which is commonly known as the frame of discernment, the power set 2^Θ is the set of all subsets of Θ , including itself and the null set φ . The basic probability assignment (BPA), $m(A)$, also known as the mass function, is assigned to an individual in a subset of the power set. The BPA, $m(A)$, is defined on the bounded interval $[0,1]$, and it satisfies

$$\begin{cases} m(\varphi) = 0, \\ \sum_{A \subseteq \Theta} m(A) = 1. \end{cases} \tag{6}$$

The belief function (Bel) is a function of each subset A of Θ , and $Bel(A)$ is a measure of one’s belief that proposition A is correct. It is defined as follows:

$$Bel(A) = \sum_{B \subset A} m(B) (\forall A \subset \Theta). \tag{7}$$

Dempster’s rule of combination, which provides an approach through which several belief functions can be determined in the same Θ , is defined as

$$m_{joint}(C) = (m_1 \oplus m_2)(C) = \begin{cases} \frac{1}{1-K} \sum_{A_1 \cap A_2 = C} m_1(A_1)m_2(A_2), & \forall C \subset \Theta, C \neq \varphi, \\ 0, & C = \varphi, \end{cases} \tag{8}$$

where K is the normalization factor,

$$K = \sum_{A_1 \cap A_2 = \varphi} m_1(A_1)m_2(A_2) < 1, \tag{9}$$

which denotes the degree of conflict between A_1 and A_2 , with 0 representing no conflict and 1 representing total conflict. If K is too large, Dempster’s rule of combination will fail. With Eqs. (8) and (9), the fusion vector R is

$$R = [r_1, r_2, \dots, r_N], \tag{10}$$

where $r_i (i = 1, 2, \dots, N)$ represents the output of each class in the frame of discernment, Θ . To assist in numeric decision-making for weld defect classification, the “decision rule” is formulated as follows: if $r_i > \eta (i = 1, 2, \dots, N)$ and r_i is the maximum, then target S is of class X_i . In this study, the threshold η is set as 0.50.

The following example serves to describe the detailed steps of the calculation of R . Supposing a frame of discernment $\Theta = \{A_1, A_2, \theta\}$ where each proposition in Θ symbolically represents two weld defects classes (i.e., A_1 represents PO and A_2 represents SL), and θ represents the uncertainty of a defect belonging to A_i , three BPAs (m_1, m_2 , and m_3) are considered in this frame of discernment:

$$\begin{aligned} m_1 < A_1, A_2, \theta > &= m_1 < 0.8, 0.1, 0.1 >, \\ m_2 < A_1, A_2, \theta > &= m_2 < 0.6, 0.3, 0.1 >, \\ m_3 < A_1, A_2, \theta > &= m_3 < 0.75, 0.2, 0.05 >. \end{aligned}$$

Then, the new BPA, m_{joint} , is calculated based on m_1 and m_2 with Eqs. (8) and (9):

$$\begin{aligned} K &= \sum_{A_1 \cap A_2 = \varphi} m_1(A_1)m_2(A_2) \\ &= m_1(A_1)m_2(A_2) + m_1(A_2)m_2(A_1) = 0.24 + 0.06 = 0.3, \\ m_{joint}(A_1) &= \frac{m_1(A_1)m_2(A_1) + m_1(A_1)m_2(\theta) + m_1(\theta)m_2(A_1)}{1 - K} \\ &= \frac{0.48 + 0.08 + 0.06}{1 - 0.3} = 0.8857, \\ m_{joint}(A_2) &= \frac{m_1(A_2)m_2(A_2) + m_1(A_2)m_2(\theta) + m_1(\theta)m_2(A_2)}{1 - K} \\ &= \frac{0.03 + 0.03 + 0.01}{1 - 0.3} = 0.1, \\ m(\theta) &= \frac{m_1(\theta)m_2(\theta)}{1 - K} = \frac{0.01}{1 - 0.3} = 0.0143. \end{aligned}$$

In the same manner, m_{joint} can be combined with m_3 as:

$$\begin{aligned} K &= m_{joint}(A_1)m_3(A_2) + m_{joint}(A_2)m_3(A_1) \\ &= 0.11714 + 0.075 = 0.25214, \\ m'_{joint}(A_1) &= \frac{m_{joint}(A_1)m_3(A_1) + m_{joint}(A_1)m_3(\theta) + m(\theta)m_3(A_1)}{1 - K} \\ &= 0.9618, \\ m'_{joint}(A_2) &= 0.0373, \\ m'(\theta) &= \frac{m(\theta)m_3(\theta)}{1 - K} = \frac{0.0143 \times 0.05}{1 - 0.25214} = 0.0009. \end{aligned}$$

Here, the fusion vector, R , is defined as

$$R = [r_1, r_2, r_3] = [m'_{\text{joint}}(A_1), m'_{\text{joint}}(A_2), m'(\theta)] = [0.9618, 0.0373, 0.0009].$$

Finally, each r_j ($j = 1, 2, 3$) represents the output of each class in the frame of discernment $\Theta = \{A_1, A_2, \theta\}$, and r_j represents the belief committed exactly to A_j . This example shows that the weld defect class is most likely class A_1 (porosity) because r_1 is the maximum value in the fusion vector, R .

SVF identification in frame of discernment

Given N classes of weld defects and M features, the weld defect class can be described as a standard class matrix, S_0 :

$$S_0 = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1M} \\ x_{21} & x_{22} & \cdots & x_{2M} \\ \vdots & \vdots & \cdots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NM} \end{bmatrix}. \tag{11}$$

In addition, the frame of discernment can be defined as $\Theta = \{X_1, X_2, \dots, X_N\}$, where X_i represents the i -th class of weld defect.

When using the DS method, the first problem is deciding Θ by determining the SVF, x_{ij} ($j = 1, 2, \dots, M$), for X_i . Because the x_{ij} have different values, some will be too large, while others will be small, so it is necessary to develop a reliable method for obtaining a standard x_{ij} . Generally, the value of feature x_{ij} approximately obeys the normal distribution, so a method was developed to calculate feature x_{ij} based on FHA (Dudewicz 1999). As shown in Fig. 3, FHA is a statistical tool for presenting numerous data in a form that clarifies the central tendency and the dispersion along the scale of measurement. Many feature samples, x_{ij} , were used to construct the histograms, wherein each column represents the frequency of a certain value of x_{ij} . Then, the standard deviation of x_{ij} , which follows a normal probability distribution (NPD), is calculated and used as an estimate of σ . Finally, because $\pm 3\sigma$ standard deviations includes 99.73% of the population, most x_{ij} are included for calculating the SVF, x_{ij} , and the mean value of samples (MVS) can be calculated if the samples belong to $\pm 3\sigma$ standard deviations; otherwise, if samples are beyond the $\pm 3\sigma$ standard deviations, they should not be used for the calculation. Then, the MVS can be used as the SVF, x_{ij} , and the frame of discernment Θ can be obtained, which is more reliable for defect classification.

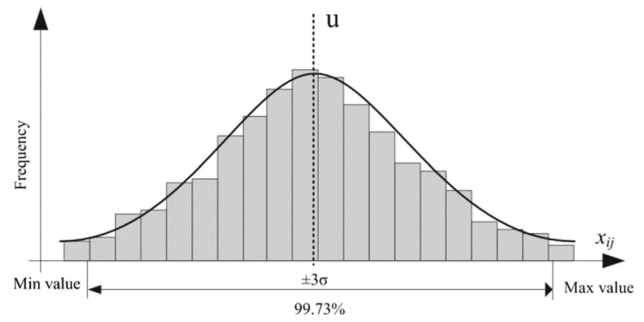


Fig. 3 Frequency histograms analysis (FHA) for feature x_{ij}

Improved Dempster’s rule of combination based on WF

As shown in “Brief introduction of DS evidence theory” section, the calculation of m is another crucial aspect that determines whether DS evidence theory will succeed. In this paper, the BPA, m , is calculated based on the WF of each feature and the frame of discernment Θ ; namely, the main steps of calculating m are similar to the method proposed in previous work (Jiang et al. 2016).

Supposing that the target weld defect, S , contains the testing data that need to be classified, it can be defined as the vector

$$S = [s_1, s_2, \dots, s_j, \dots, s_M], \tag{12}$$

where s_j represents the j -th feature of target weld defect S .

Based on Eqs.(11) and (12), the similarity matrix, P , which indicates the similarity of S to S_0 , can be obtained as

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N1} & \cdots & p_{NM} \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_N \end{bmatrix}, \tag{13}$$

where

$$p_{ij} = 1/d_{ij}, \tag{14}$$

and

$$d_{ij} = |s_j - x_{ij}| \quad i = 1, 2, \dots, N; j = 1, 2, \dots, M. \tag{15}$$

As shown in Eq.(13), a larger P_i indicates that the target weld defect, S , more likely belongs to the i -th class of weld defect. Conversely, a smaller P_i indicates that the target weld defect, S , less likely belongs to the i -th class of weld defect.

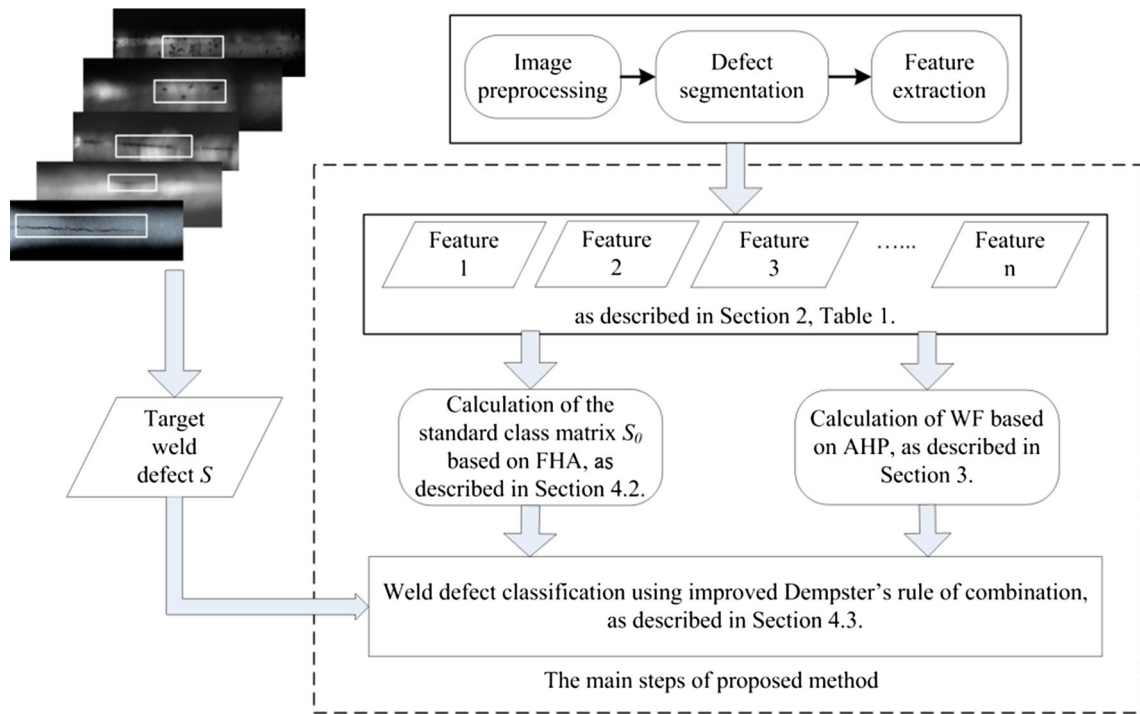


Fig. 4 Procedure for weld defect classification based on the proposed method

To obtain the BPA, m , the normalizing P is calculated via

$$p'_{ij} = p_{ji} / \sum_{k=1}^N p_{kj}, \tag{16}$$

such that

$$P' = \begin{bmatrix} p'_{11} & p'_{12} & \cdots & p'_{1N} \\ p'_{21} & p'_{22} & \cdots & p'_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p'_{M1} & p'_{M2} & \cdots & p'_{MN} \end{bmatrix} = \begin{bmatrix} P'_{m_1} \\ P'_{m_2} \\ \vdots \\ P'_{m_M} \end{bmatrix}. \tag{17}$$

Then, P' can be defined as the BPA, m , of X_j , which satisfies Eq. (6):

$$\begin{bmatrix} m_1(X_1) & m_1(X_2) & \cdots & m_1(X_N) \\ m_2(X_1) & m_2(X_2) & \cdots & m_2(X_N) \\ \vdots & \vdots & \ddots & \vdots \\ m_M(X_1) & m_M(X_2) & \cdots & m_M(X_N) \end{bmatrix} = \begin{bmatrix} P'_{m_1} \\ P'_{m_2} \\ \vdots \\ P'_{m_M} \end{bmatrix}. \tag{18}$$

Considering the WF, which was calculated in ‘‘Calculation of WF based on AHP’’ section, to account for the different WFs in defect classification, Dempster’s rule of combination was revised as follows:

$$m'_i(X) = w_i \cdot m_i(X) \quad (i = 1, 2, \dots, M). \tag{19}$$

Then, Dempster’s rule of combination from Eqs. (8) and (9) are revised as:

$$m_{joint}(C) = (m'_1 \oplus m'_2)(C) = \begin{cases} \frac{1}{1-K} \sum_{A_1 \cap A_2 = C} m'_1(A_1)m'_2(A_2) & \forall C \in \Theta, C \neq \varphi \\ 0 & C = \varphi \end{cases}, \tag{20}$$

where

$$K = \sum_{A_1 \cap A_2 = \varphi} m'_1(A_1)m'_2(A_2) < 1. \tag{21}$$

The fusion vector, R , can be calculated based on $m'_i (i = 1, 2, \dots, M)$, and the main steps of the method can be described as follows:

- (a) Combining m'_1 with m'_2 using Eqs. (20) and (21) yields m_{joint} .
- (b) Combining m_{joint} with m'_3 using Eqs. (20) and (21) yields a new m_{joint} .
- (c) Repeating step (b) for the remaining $m'_i (i = 1, 2, \dots, M)$ yields the final fusion vector R .

In summary, the main steps of the weld defects classification method based on AHP and DS evidence theory are shown in Fig. 4. Firstly, several techniques, such as image

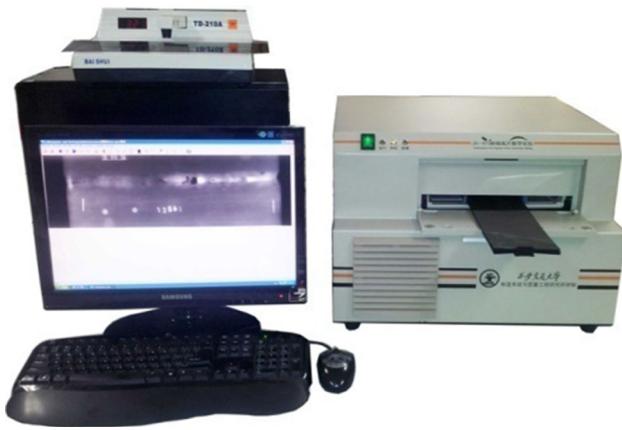


Fig. 5 Automatic radiographic testing system (JD-RTD)

preprocessing, defect segmentation, and feature extraction, are performed sequentially to obtain features for defect classification. Second, as shown in the dashed-line box in Fig. 4, the procedure of the proposed method includes three steps:

- (a) The WF of the selected features is calculated based on AHP using Eqs. (1)–(5).
- (b) The standard class matrix, S_0 , is calculated based on FHA using Eq. (11).
- (c) The weld defect is classified using an improved Dempster’s method based on the S_0 and WF of each feature. Specifically, the BPA, m , is calculated using Eqs. (12)–(18), and the fusion vector, R , is calculated using Eqs. (19)–(21).

Weld defect classification case study

To illustrate the effectiveness of the method proposed in this paper, 172 samples covering 5 classes of weld defects were selected: 32 PO, 41 SL, 38 LP, 35 LF, and 26 CR. The radiographic images were provided by the Dongfang Turbine Co., Ltd, Sichuan, China, and the samples were produced from the digitized radiographic images using an X-ray film scanner (JD-RTD) that was developed in-house, as shown in Fig. 5.

With these five classes of weld defect, the weld defects’ frame of discernment can be defined as

$$\Theta = \{X_1, X_2, X_3, X_4, X_5\}, \tag{22}$$

where X_1, X_2, X_3, X_4 , and X_5 represent PO, SL, LP, LF, and CR, respectively. First, as demonstrated in “Calculation

of WF based on AHP” section, the comparison matrix, A , can be obtained via AHP:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} & L & Ar & Sp & Re & Rr & Sk & Ku \\ L & 1 & 0.5 & 4 & 3 & 0.33 & 5 & 6 \\ Ar & 2 & 1 & 5 & 3 & 0.5 & 6 & 7 \\ Sp & 0.25 & 0.2 & 1 & 0.5 & 0.17 & 2 & 3 \\ Re & 0.5 & 0.33 & 2 & 1 & 0.2 & 3 & 5 \\ Rr & 3 & 2 & 6 & 5 & 1 & 7 & 9 \\ Sk & 0.2 & 0.17 & 0.5 & 0.33 & 0.14 & 1 & 2 \\ Ku & 0.17 & 0.14 & 0.33 & 0.2 & 0.11 & 0.5 & 1 \end{bmatrix}. \tag{23}$$

From Eqs. (2) and (3), $CR = 0.034 < 0.1$, which means that the consistency is acceptable. Therefore, the WF can be obtained via Eqs. (4) and (5):

$$W = \{w_1 \ w_2 \ w_3 \ w_4 \ w_5 \ w_6 \ w_7\}$$

$$= \{0.17 \ 0.24 \ 0.06 \ 0.1 \ 0.37 \ 0.04 \ 0.03\}. \tag{24}$$

For the frame of discernment, Θ , all 172 samples were selected for calculation of the standard class matrix, S_0 , in the frame of discernment based on the method described in “SVF identification in frame of discernment” section; the results of which are shown in Table 2.

After obtaining the standard class matrix, S_0 , one weld defect sample is selected as the target weld defect, S , and the method described in “Improved Dempster’s rule of combination based on WF” section is applied to obtain the fusion vector $R = [r_1, r_2, r_3, r_4, r_5]$, where each r_j ($j = 1, 2, \dots, 5$) represents the output of each X_i in the frame of discernment $\Theta = \{X_1, X_2, X_3, X_4, X_5\}$, respectively. For each case analysis, each correct recognition rate ($CRR(i)(i = 1, 2, 3)$) is determined by the maximum, second-maximum, and third-maximum of R , respectively.

The following example illustrates the calculation of the CRR. Four samples of defect type X_1 (PO) are selected as the target weld defects, S , and the result, R , is shown in Table 3, in which the bold entries indicate the maximum values of R_i .

From Table 3, taking sample no. 1 as S , the fusion vector is $R = [r_1, r_2, r_3, r_4, r_5] = [0.93987, 0.00448, 0.00206, 0.03830, 0.01532]$; thus, $r_1 = 0.93987$ satisfies the decision rule, as $r_1 > 0.5$ and it is the maximum of the fusion vector, so the target weld defect, S , can be correctly classified as X_1 (PO). Because there is only one target weld defect S for which r_1 is the maximum, the $CRR(1)$ for classification is:

$$CRR(1) = 1/4 = 25\%.$$

Table 2 Standard class matrix, S_0

S_0	L	Ar	Sp	Re	Rr	Sk	Ku
X_1	11.1200	1.0563	0.3452	1.7240	186.8909	0.0319	0.6450
X_2	17.1227	2.3644	0.3907	1.2574	93.7009	0.0680	0.6417
X_3	24.0450	13.0200	0.4943	0.3455	44.8030	0.0710	0.6584
X_4	28.3810	5.6470	0.4310	0.9209	79.2498	0.0308	0.6406
X_5	8.2130	22.1984	0.5010	0.3637	160.5996	0.0309	0.6510

Considering sample no. 2 of S , $r_1 = 0.19593$ and $r_2 = 0.69580$; this is a classification error in that S is classified as SL with the decision rule. Moreover, because $r_2 > r_1 > r_5 > r_4 > r_3$, S is less likely to be a PO weld defect than an SL weld defect. That is, if the classification is decided by the second-maximum r_1 , the result is correct. Therefore, the r_i can be arranged in descending order, representing the corresponding possibility that S is classified as X_i . Because r_1 (0.19593) is the second-maximum,

$$CRR(2) = 1/4 + CRR(1) = 50\%.$$

Considering sample no. 3 of S , CRR(3) can similarly be calculated:

$$CRR(3) = 1/4 + CRR(2) = 75\%.$$

Results

The resulting classifications of produced by the proposed method are shown in Table 4, which includes 172 samples covering 5 classes of weld defects. For LF and CR weld defects, the proposed method has excellent performance with CRR(1) at 91.42 and 96.15%, respectively. Because LF and CR weld defects differ greatly from the other weld defect classes, they can easily be classified correctly by using the WF. Because SL and LP weld defects are respectively similar to PO and CR weld defects, the method has poor performance with CRR(1) at 68.29 and 73.68% for SL and LP weld defects, respectively. Ultimately, for PO, SL, and LP, the proposed method has good performance with CRR(3) at 100%. This indicates that the method proposed in this paper yields significant separability, so it is suitable for classification of weld defects.

Discussion

Table 5 shows the aggregate CRR of the proposed method for all test samples; additionally, the performance of the proposed method is compared to those of the traditional DS method [7, 27, DMSVM, PCA-SVM, and MLP-ANN. The proposed method is more accurate than the traditional DS

Table 3 Examples of R

Sample no.	Fusion vector R of S				
	r_1	r_2	r_3	r_4	r_5
1	0.93987	0.00448	0.00206	0.03830	0.01532
2	0.19539	0.69580	0.00126	0.04245	0.06511
3	0.02901	0.87374	0.00451	0.08049	0.01226
4	0.00029	0.97915	0.01869	0.00172	0.00017

method because the CRR(1) and CRR(2) values obtained using the former method are higher than those from the latter method. This result indicates that the technique combined with the calculation of WF based on AHP strongly influences the results of the proposed method, i.e., reasonable weights are assigned to different features through AHP and the proposed technique improves the classification accuracy. For example, as for the LF and CR weld defects, the features of W, Ar, and Rr are very important factors in weld defect classification. Therefore, the CRR is higher when the WF of W, Ar, and Rr are higher. Meanwhile, because FHA can capture the feature’s distribution, the S_0 calculated based on FHA with few training samples can be more reliable for defect classification, improving the performance of the proposed method.

Although the CRR(1) of the proposed method is only 81.98%, CRR(2) is larger than the CRRs of DMSVM, PCA-SVM and MLP-ANN. If the three largest values of r_i are used together to decide the classification of the weld defect, CRR(3) is 100%. Additionally, an engineer sometimes needs to instead be given the probability of a defect belonging to the different categories—this accounts for some weld defects that cannot be classified due to the similarity between different types of defects, e.g., between LP and CR or between PO and LF. Thus, the proposed method not only has good classification performance, but it can also address the uncertainty associated with weld defect classes, i.e., by calculating the combinations of mass functions, the proposed method determines the belief committed to a weld defect class. Additionally, because the proposed method is an information fusion technique, it can be applied to some pattern recognition problems, e.g., fault diagnosis, target recognition and disease diagnosis.

Table 4 Classification results of weld defects

Target class (Actual #)	Correct recognition rate of classification		
	CRR(1) (%)	CRR(1,2) (%)	CRR(1,2,3) (%)
PO (32)	87.5	90.63	100
SL (41)	68.29	87.80	100
LP (38)	73.68	89.47	100
LF (35)	91.43	100	–
CR (26)	96.15	100	–

Table 5 Comparison of performance

Method	Correct recognition rate of classification		
	CRR(1) (%)	CRR(2) (%)	CRR(3) (%)
Proposed method	81.98	93.02	100
Traditional DS method	78.49	88.37	100
DMSVM	88.95	–	–
PCA-SVM	81.97	–	–
MLP-ANN	87.21	–	–

However, the classification method based on DS also has some shortcomings. The first is deciding the frame of discernment Θ , i.e., deciding the SVF in the frame of discernment. Generally, the SVF can greatly improve the separability of elements in the frame of discernment. The second one is calculating the BPA, m , which is important for achieving higher defect classification accuracy. Additionally, as shown in “Brief introduction of DS evidence theory” section, accurate determination of m is essential for DS evidence theory to succeed. Furthermore, each method—including DMSVM, PCA-SVM, MLP-ANN, and the proposed method—has its limitations, e.g., PO, SL, and LP could each only be classified by some of the methods. Using two or more complementary methods would be a good solution to this problem.

Conclusion

In this paper, a method for classifying weld defects based on AHP and DS evidence theory was proposed; this new method mainly includes three novel contributions: (1) calculation of the WF of a weld defect based on AHP, (2) determination of the SVF in the frame of discernment, and (3) an improved Dempster’s rule for combination. A case study of weld defect classification was provided to illustrate the method and test its effectiveness. The results show that the proposed method can fuse different features to effectively make classification decisions based on few training samples and that it has better performance than traditional DS methods. Additionally, although the CRR(1) of the proposed method is slightly smaller than those of traditional methods, such as DMSVM and PCA-SVM, its CRR(2) is significantly larger than the CRRs of traditional methods. Notably, the proposed method

can address the uncertainty associated with weld defect classification, i.e., it can give the probability of a defect belonging to different defect types if there is little difference between the weld defects.

Future work will focus on (1) studying the method for weld defect feature extraction and (2) calculation of the mass function and combined rules of DS evidence theory. In this paper, only seven features were used for classification; the CRR can be significantly improved if additional suitable features can be selected. Additionally, the mass function, also called the BPA, is crucial for successfully applying DS evidence theory, so better combination rules could eliminate conflicts between various forms of evidence. In fact, the classification error in the proposed method mostly resulted from conflicts between weld features. Therefore, the mass function and combination rules should be investigated in future studies.

Acknowledgements The authors sincerely thank the referees for their helpful suggestions and comments, which greatly improved the quality of the paper. This research was supported by the National Natural Science Foundation of China (Grant No. 51375375).

References

- da Silva, R. R., & Mery, D. (2007a). The state of the art of weld seam radiographic testing: Part I—image processing. *Materials Evaluation*, 65(6), 643–647.
- da Silva, R. R., & Mery, D. (2007b). The state of the art of weld seam radiographic testing: Part II—pattern recognition. *Materials Evaluation*, 65(9), 833–838.
- Dempster, A. P. (1968). A generalization of Bayesian inference. *Journal of the Royal Statistical Society: Series B (Methodology)*, 30(2), 205–247.
- Dudewicz, E. J. (1999). Basic statistical methods. In J. M. Juran & A. B. Godfrey (Eds.), *Juran’s quality handbook* (5th ed., pp. 44.1–44.112). New York: McGraw-Hill.

- Du, X., Shen, Y., & Wang, Y. (2008). Weld defect classification in ultrasonic testing basing on time-frequency discriminant features. *Transactions-China Welding Institution*, 29(2), 89–92.
- Fan, M., Wei, L., He, Z., Wei, W., & Lu, X. (2016). Defect inspection of solder bumps using the scanning acoustic microscopy and fuzzy SVM algorithm. *Microelectronics Reliability*, 65, 192–197. <https://doi.org/10.1016/j.microrel.2016.08.010>.
- Gao, H., Shen, X., Jiang, Z., Yang, H., & Yan, L. (2012). Image subcategory classification based on Dempster–Shafer evidence theory. In *International Conference on Computer Science and Service System* (pp. 2289–2292). Nanjing: CHN, August 11–13, 2012. <https://doi.org/10.1109/CSSS.2012.568>
- Gao, W., & Hu, Y. H. (2014). Real-time X-ray radiography for defect detection in submerged arc welding and segmentation using sparse signal representation. *Insight-Non-Destructive Testing and Condition Monitoring*, 56(6), 299–307. <https://doi.org/10.1784/insi.2014.56.6.299>.
- Gu, K., Zhai, G., Yang, X., & Zhang, W. (2013). A new reduced-reference image quality assessment using structural degradation model. In *2013 IEEE international symposium on circuits and systems (ISCAS)*. (pp. 1095–1098). Beijing: CHN, May 19–23, 2013. <https://doi.org/10.1109/ISCAS.2013.6572041>.
- Hafizan, C., Noor, Z. Z., Abba, A. H., & Hussein, N. (2016). An alternative aggregation method for a life cycle impact assessment using an analytical hierarchy process. *Journal of Cleaner Production*, 112(4), 3244–3255. <https://doi.org/10.1016/j.jclepro.2015.09.140>.
- Han, D., Yang, Y., & Han, C. (2014). Advances in DS evidence theory and related discussions. *Control and Decision*, 29(1), 1–11. <https://doi.org/10.13195/j.kzyjc.2013.0517>.
- Hong, S. J., Lim, W. Y., Cheong, T., & May, G. S. (2012). Fault detection and classification in plasma etch equipment for semiconductor manufacturing e-diagnostics. *IEEE Transactions on Semiconductor Manufacturing*, 25(1), 83–93. <https://doi.org/10.1109/TSM.2011.2175394>.
- Jiang, H., Liang, Z., Gao, J., & Dang, C. (2016). Classification of weld defect based on information fusion technology for radiographic testing system. *Review of Scientific Instruments*, 87(3), 035110. <https://doi.org/10.1063/1.4943220>.
- Li, T.-S. (2009). Applying wavelets transform, rough set theory and support vector machine for copper clad laminate defects classification. *Expert systems with Applications*, 36(3 Pt 2), 5822–5829. <https://doi.org/10.1016/j.eswa.2008.07.040>.
- Liao, T. W. (2003). Classification of welding flaw types with fuzzy expert systems. *Expert Systems with Applications*, 25(1), 101–111. [https://doi.org/10.1016/S0957-4174\(03\)00010-1](https://doi.org/10.1016/S0957-4174(03)00010-1).
- Liao, T. W. (2009). Improving the accuracy of computer-aided radiographic weld inspection by feature selection. *NDT & E International*, 42(4), 229–239. <https://doi.org/10.1016/j.ndteint.2008.11.002>.
- Lim, T. Y., Ratnam, M. M., & Khalid, M. A. (2007). Automatic classification of weld defects using simulated data and an MLP neural network. *Insight-Non-Destructive Testing and Condition Monitoring*, 49(3), 154–159. <https://doi.org/10.1784/insi.2007.49.3.154>.
- Liu, H., Chen, Y., Peng, X., & Xie, J. (2011). A classification method of glass defect based on multiresolution and information fusion. *The International Journal of Advanced Manufacturing Technology*, 56(9–12), 1079–1090. <https://doi.org/10.1007/s00170-011-3248-z>.
- Maruthur, N. M., Joy, S., Dolan, J., Segal, J. B., Shihab, H. M., & Singh, S. (2013). Systematic assessment of benefits and risks: Study protocol for a multi-criteria decision analysis using the analytic hierarchy process for comparative effectiveness research. *F1000Research*, 2, 160. <https://doi.org/10.12688/f1000research.2-160.v1>
- Movafeghi, A. (2015). Using empirical mode decomposition and a fuzzy algorithm for the analysis of weld defect images. *Insight-Non-Destructive Testing and Condition Monitoring*, 57(1), 35–39. <https://doi.org/10.1784/insi.2014.57.1.35>.
- Mu, W., Gao, J., Jiang, H., Wang, Z., Chen, F., & Dang, C. (2013). Automatic classification of weld defects based on optimal PCA and SVM. *Insight-Non-Destructive Testing and Condition Monitoring*, 55(10), 535–539. <https://doi.org/10.1784/insi.2012.55.10.535>.
- Mu, W., Gao, J., Wang, Z., Jiang, H., Chen, F., & Dang, C. (2013). Radiographic image assessment approach based on human visual system. *Journal of Xi'an Jiaotong University*, 47(7), 91–95.
- Nacereddine, N., Hamami, L., & Ziou, D. (2006). Thresholding techniques and their performance evaluation for weld defect detection in radiographic testing. *International Journal of Machine Graphics and Vision*, 15(3), 557–566.
- Nacereddine, N., Ziou, D., & Hamami, L. (2013). Fusion-based shape descriptor for weld defect radiographic image retrieval. *The International Journal of Advanced Manufacturing Technology*, 68(9–12), 2815–2832. <https://doi.org/10.1007/s00170-013-4857-5>.
- Pan, J., Jiang, H., Gao, J., & Yang, P. (2011). Condition diagnosis with complex network-time series analysis. In *Proceedings of Annual Reliability and Maintainability Symposium, Lake Buena Vista, FL, USA* (pp. 1–6), January 24–27, 2011. <https://doi.org/10.1109/RAMS.2011.5754502>.
- Pohl, C., & Van Genderen, J. L. (1998). Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, 19(5), 823–854. <https://doi.org/10.1080/014311698215748>.
- Saaty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.
- Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, 48(1), 9–26. [https://doi.org/10.1016/0377-2217\(90\)90057-1](https://doi.org/10.1016/0377-2217(90)90057-1).
- Salchak, Y., Tverdokhlebova, T., Sharavina, S., & Lider, A. (2016). The classification of weld seam defects for quantitative analysis by means of ultrasonic testing. *IOP Conference Series: Materials Science and Engineering*, 132, 012027. <https://doi.org/10.1088/1757-899X/132/1/012027>.
- Shafer, G. (1976). *A mathematical theory of evidence*. Princeton: Princeton University Press.
- Shen, Q., & Gao, J. (2010). Improving the classification accuracy of the weld defect by chaos-search-based feature selection. *Insight-Non-Destructive Testing and Condition Monitoring*, 52(10), 530–539. <https://doi.org/10.1784/insi.2010.52.10.530>.
- Shen, Q., Gao, J., & Li, C. (2010). Automatic classification of weld defects in radiographic images. *Insight-Non-Destructive Testing and Condition Monitoring*, 52(3), 1–6. <https://doi.org/10.1784/insi.2010.52.3.134>.
- Sreedhar, U., Krishnamurthy, C. V., Balasubramaniam, K., Raghupathy, V. D., & Ravisankar, S. (2012). Automatic defect identification using thermal image analysis for online weld quality monitoring. *Journal of Materials Processing Technology*, 212(7), 1557–1566. <https://doi.org/10.1016/j.jmatprotec.2012.03.002>.
- Wang, G., & Liao, T. W. (2002). Automatic identification of different types of welding defects in radiographic images. *NDT & E International*, 35(8), 519–528. [https://doi.org/10.1016/S0963-8695\(02\)00025-7](https://doi.org/10.1016/S0963-8695(02)00025-7).
- You, D., Gao, X., & Katayama, S. (2015). WPD-PCA-based laser welding process monitoring and defects diagnosis by using FNN and SVM. *IEEE Transactions on Industrial Electronics*, 62(1), 628–636. <https://doi.org/10.1109/TIE.2014.2319216>.
- Zahran, O., & Al-Nuaimy, W. (2002). Recent developments in ultrasonic techniques for rail-track inspection. In *Proceedings of the Annual Conference of the British Institute of Non-destructive Testing (BINDT 2002)* (pp. 55–60). Southport: GBR, September 17–19, 2002.

- Zahran, O., Kasban, H., El-Kordy, M., & Abd El-Samie, F. E. (2013). Automatic weld defect identification from radiographic images. *NDT & E International*, 57, 26–35. <https://doi.org/10.1016/j.ndteint.2012.11.005>.
- Zapata, J., Vilar, R., & Ruiz, R. (2010). An adaptive-network-based fuzzy inference system for classification of welding defects. *NDT & E International*, 43(3), 191–199. <https://doi.org/10.1016/j.ndteint.2009.11.002>.
- Zapata, J., Vilar, R., & Ruiz, R. (2012). Automatic inspection system of welding radiographic images based on ANN under a regularisation process. *Journal of Nondestructive Evaluation*, 31(1), 34–45. <https://doi.org/10.1007/s10921-011-0118-4>.
- Zhang, X., Zhu, Z., Xu, J., & Ren, S. (2005). The classification algorithm of defects in weld image based on asymmetrical SVMs. In *International Conference on Control Automation 2005 (ICCA '05)* (pp. 1215–1219). Budapest: HUN, June 26–29, 2005. <https://doi.org/10.1109/ICCA.2005.1528306>
- Zhu, P., Yin, C., Cheng, Y., Huang, X., Cao, J., Vong, C.-M., et al. (2017). An improved feature extraction algorithm for automatic defect identification based on eddy current pulsed thermography. *Mechanical Systems and Signal Processing*. <https://doi.org/10.1016/j.ymsp.2017.02.045>.