

# A RECOGNITION ALGORITHM TO DETECT PIPE WELD DEFECTS

*Wei Cui, Ke Wang, Qiang Zhang, Peng Zhang*

Preliminary communication

Taking magnetic flux leakage (MFL) imaging of pipe weld defects as the research object, a weld defect image recognition algorithm based on grey-gradient co-occurrence matrix (GGCM) and cluster analysis and mathematical morphology is proposed. Recognition of different types of welding defects was achieved. Firstly, a continuous non-contact scanning MFL system for the pipe weld was used to collect the three-dimensional MFL. Secondly, the three-dimensional MFL signal was converted to a two-dimensional grayscale image. Then the MFL image characteristics of the two-dimensional grayscale image were extracted using GGCM. Based on extracted image features, the characteristic quantity was analysed by using  $k$ -means clustering and then through the combination of histogram equalization, Otsu's method of binaryzation, morphologically removing small objects, edge detection, and then structuring a morphologically optimized edge extraction method for edge detection on the grayscale. Through combination of several methods, a new algorithm to improve the detection effect was structured. The results indicated that this algorithm is adaptable and practical. This algorithm solved difficulties associated with the MFL method being used in the weld testing to realize the recognition of pipe weld defects and break through the applicable limitations of traditional signal processing technology.

**Keywords:** defects; magnetic flux leakage imaging; pipe weld; recognition algorithm

## Algoritam raspoznavanja za otkrivanje grešaka u zavarima cijevi

Prethodno priopćenje

Uzimajući kao objekt istraživanja propuštanje magnetskog toka (MFL) za snimanje grešaka u zavarima cijevi, predložen je algoritam raspoznavanja slika grešaka u zavarima utemeljen na matrici istovremene pojave sivih tonova (GGCM) te klsterskoj analizi i matematičkoj morfologiji. Postignuto je raspoznavanje različitih vrsta grešaka u zavarima. Prvo je korišten sustav kontinuiranog bezdodirnog skeniranja MFL za prikupljanje trodimenzijskog propuštanja magnetskog toka (MFL), a nakon toga je trodimenzijski MFL signal pretvoren u dvodimenzijsku sliku sivih tonova. Zatim su karakteristike MFL slike za dvodimenzijsku sliku u sivim tonovima izdvojene pomoću GGCM. Na temelju izdvojenih značajki slike, analizira se karakteristična količina pomoću particioniranja  $k$ -sredine, a zatim kroz kombinaciju izjednačenja histograma, Otsuovu metoda binarizacija, morfološkog uklanjanja malih objekata, otkrivanja rubova, a zatim strukturiranja morfološki optimiziranog algoritma ekstrakcije ruba za otkrivanje rubova na sivim tonovima. Kombiniranjem nekoliko metoda, strukturira se novi algoritam za poboljšanje učinka otkrivanja. Rezultati su pokazali da je ova metoda prilagodljiva i praktična. Ovaj je algoritam riješio poteškoće vezane uz MFL metodu koja se koristi u ispitivanju zavara radi otkrivanja grešaka u zavarima cijevi, a nadilazi granice primjene tradicionalne tehnologije obrade signala.

**KLjučne riječi:** algoritam raspoznavanja; greške (oštećenja); snimanje propuštanja magnetskog toka; zavari cijevi

## 1 Introduction

The weld is an important connection part of the pipe but during the process of manufacturing and through its use, it produces defects. The weld defects usually occur in the weld fusion zone along the border of the heat affected zone and the weld. Because these areas have a geometric discontinuity and similar mechanical properties and can have a concentration of stress, the area under the joint action of medium pressure is susceptible to stress, corrosion cracking, and induced cracks. Therefore, carrying out efficient detection of the pipe weld defects to ensure the normal use of the pipe has safety and economic significance.

At present, methods such as radiographic testing, ultrasonic testing, and magnetic memory testing [1-4] are used for online inspection of the weld. Efficient feature selection for neural network based detection of flaws in steel welded joints using ultrasound testing in literature [1]. Literature [2] described that Nondestructive evaluation (NDE) techniques of phased array ultrasonic testing (PAUT) and digital X-ray radiography were employed on friction stir (FS)-welded Aluminum Alloy (AA)-2219-T87 specimens. In literature [3], an analysis of the crack distribution in the weld sample is conducted from the acquired 2D radiographs and its corresponding 3D volumetric reconstruction achieved by linear coplanar digital laminography. This in situ approach opens new possibilities in the field of hot crack research by having

direct information of both the crack initiation and growth and its correlation to the welding parameters. Literature [4] shows that the parameters of the welding defects are significantly different from parameters of the defect-free regions, which can accurately identify the location of welding defects. Whether it is a welding defect, or welding residual stress, it can lead to changes of magnetic memory signal, but there is essential distinction between above two. Research on the magnetic flux leakage (MFL) method for testing weld defects has shown that this method is easy to carry out and reports demonstrate that MFL detection of weld defects is technically feasible but there is a lack of weld detection instruments and recognition of quantitative analysis of the weld defects in literature [5] and literature [6]. The one key technical problem associated with the MFL method for testing for the weld defects is when using MFL in the detection of the weld defect, the MFL signal generated by the weld contour is superimposed on the defect signal. These technical problems lead to limitations in the traditional signal processing for the detection and visual identification of the weld defects. This paper provided solutions to the problem by a new recognition algorithm based on gray-gradient co-occurrence matrix (GGCM), cluster analysis and mathematical morphology. This algorithm combines several methods to improve the detection effect. For example, in evaluation of Literature [7], the combination of the four methods yields 94% correct detection.

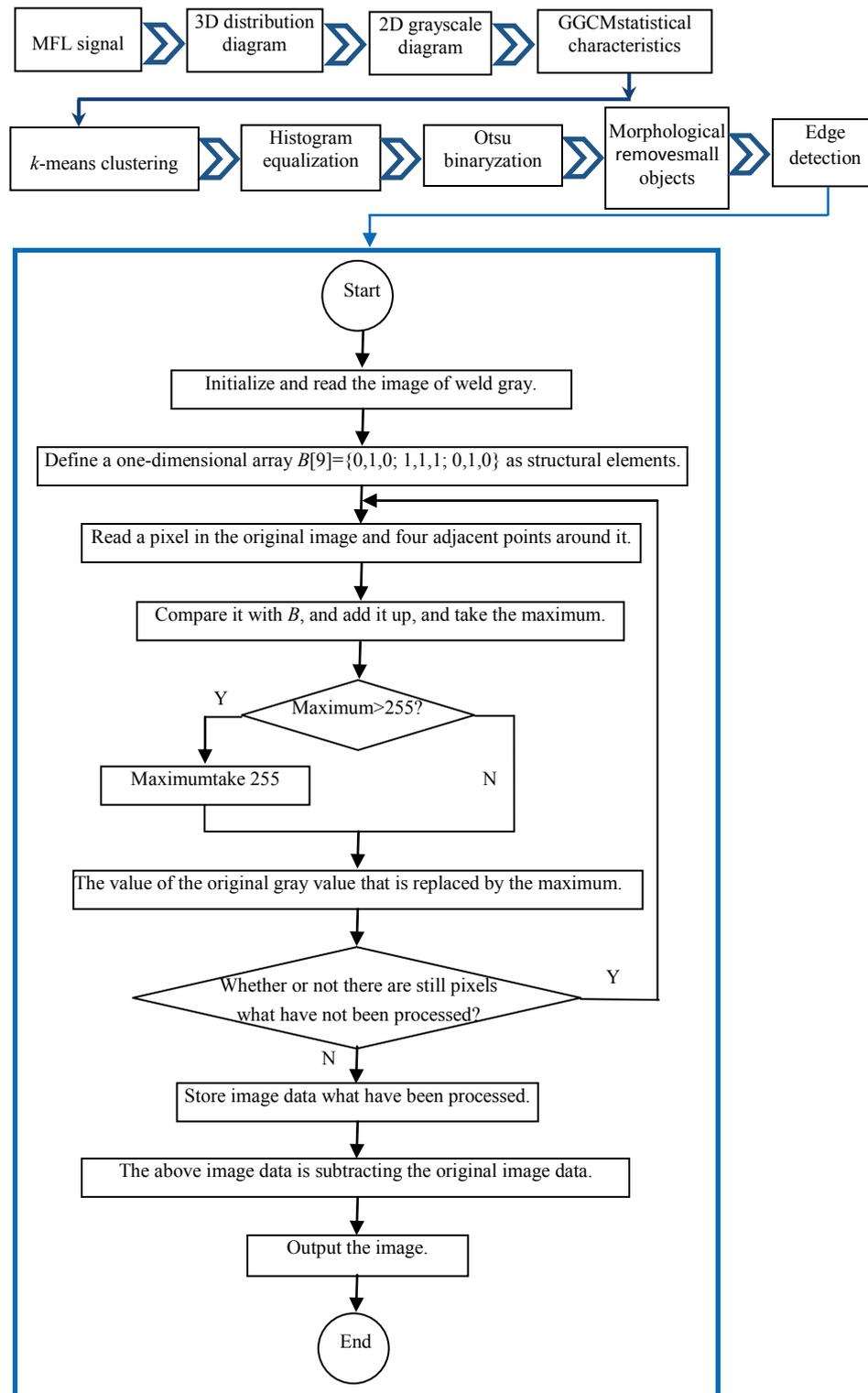


Figure 1 Recognition algorithm

## 2 Algorithm

Literature [8-11] described that the GGCM method can clearly describe the distribution of each pixel grey level and gradient in the graphics, and can also describe the space relationship of the pixel and its neighbouring pixels. This method can describe the graphics of the grayscale, shape and texture features well. GGCM can calculate all kinds of statistics of image characteristics. Characteristic parameters are small gradient advantage

characteristic ( $T_1$ ), large gradient advantage characteristic ( $T_2$ ), characteristic quantity of the gray distribution non-uniformity ( $T_3$ ), characteristic quantity of the gradient distribution non-uniformity ( $T_4$ ), energy characteristic quantity ( $T_5$ ), characteristic quantity of the gray average ( $T_6$ ), characteristic quantity of the gradient average ( $T_7$ ), characteristic quantity of the gray standard deviation ( $T_8$ ), characteristic quantity of the gradient standard deviation ( $T_9$ ), characteristic quantity of the correlation ( $T_{10}$ ), characteristic quantity of the gray entropy ( $T_{11}$ ),

characteristic quantity of the gradient entropy ( $T_{12}$ ), characteristic quantity of the hybrid entropy ( $T_{13}$ ), characteristic quantity of the inertia entropy ( $T_{14}$ ), characteristic quantity of the inverse different moment ( $T_{15}$ ). Literature [12-15] indicated that the mathematical morphology is the new theory and the new method for digital image processing and recognition, and a complete set of system theory, method and algorithm has been developed to describe the basic characteristics of images, while the traditional numerical modelling and analysis, based on the set to analyse signals has been abandoned. An edge detection operator based on mathematical morphology is a kind of nonlinear differential operator.

Because cracks are the main defects that impact the weld's mechanical properties, in many weld defects, precast crack defects on the experimental pipe and analyse distribution characteristics of the cracks in different areas (in the weld and in the fusion zone). Therefore, the algorithm in Fig. 1 is applied to identify the cracks in different positions.

Steps of the algorithm proposed in this study are:

(1) Collect three-dimensional distribution signals by MFL system for the pipe weld.

(2) Through three-dimensional distribution signals from three kinds of weld state conversion into two-dimensional grayscale.

With the distance as the horizontal axis, the number of channels as the vertical axis, the gray value of each pixel as the amplitude of the leakage magnetic signal, turn the three-dimensional magnetic flux leakage signal into two-dimensional grey image. After the transformation of grayscale patterns, the grey levels were 256, the white grey level was 255, and the black grey level was 0.

(3) Extraction of the GGCM characteristics.

(4)  $k$ -means clustering.

(5) Histogram equalization for the image.

(6) Binarization threshold was selected by using Otsu's method to process the image for binarization.

The grey value which was less than the threshold value was set at 0, and the grey value bigger than the threshold value, was set as 255.

(7) To make the image clearer, mathematical morphology was used to remove small objects.

(8) Expanding edge detection for the image by using  $E_1$  operator.

### 3 Numerical example

#### 3.1 Acquisition of three-dimensional signal

The magnetization direction is parallel to the travelling direction in the traditional detection system, but due to different characteristics of the weld structure, the traditional detection system model is no longer applicable. Therefore, this paper discussed a new detection system in which the travelling direction of the MFL detector is rotated  $90^\circ$  to make it vertical to the travelling direction. The proposed MFL testing system is similar to a scan along the longitudinal direction to detect longitudinal defects in the pipelines. Due to the change in the direction of the magnetization relative to the travelling direction, the Hall element array designed to test the new system also changes accordingly, the schematic diagram of which is shown in Fig. 2. And the continuous non-contact

scanning MFL system for the pipe weld is shown in Fig. 3.

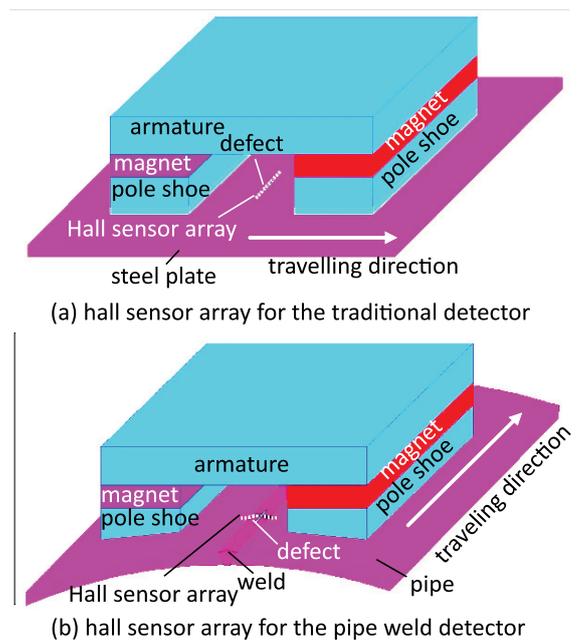


Figure 2 Diagram of Hall sensor array



(a) cracks in the weld (b) cracks in the fusion zone

Figure 3 MFL system for the pipe weld

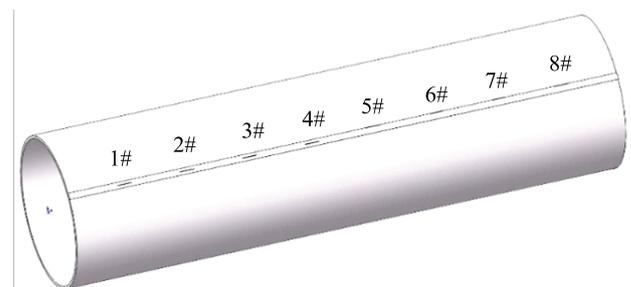


Figure 4 Diagram of the defect distribution of the Experimental board

The pipe diameter was 400 mm, and thickness was 8 mm. The weld reinforcement was 3 mm and the weld width was 18 mm. The crack length  $\times$  width = 40 mm  $\times$  1 mm. The defect depth of each group was respectively achieved at 20%, 40%, 60%, and 80% of the thickness of the pipe. The depth of 1#–4# defects (in the weld) was respectively 4.6, 6.2, 7.8, and 9.4 mm (calculated by the sum of the weld reinforcement of 3 mm and achieved 20%, 40%, 60%, and 80% thickness of the pipe, respectively). The depth of 5#–8# defects (in the fusion zone) was 1.6, 3.2, 4.8, and 6.4 mm (with 20%, 40%, 60%, and 80% thickness of the pipe, respectively). The precast cracks in the pipe are shown in Fig. 4.

Six groups for each type of defect signals were collected in the experiment process, 24 samples of each type of defect were collected, and 24 samples of the weld

without defects were also collected. A group of three-dimensional distribution signals from three kinds of weld states in the experiment are respectively shown in Fig. 5.

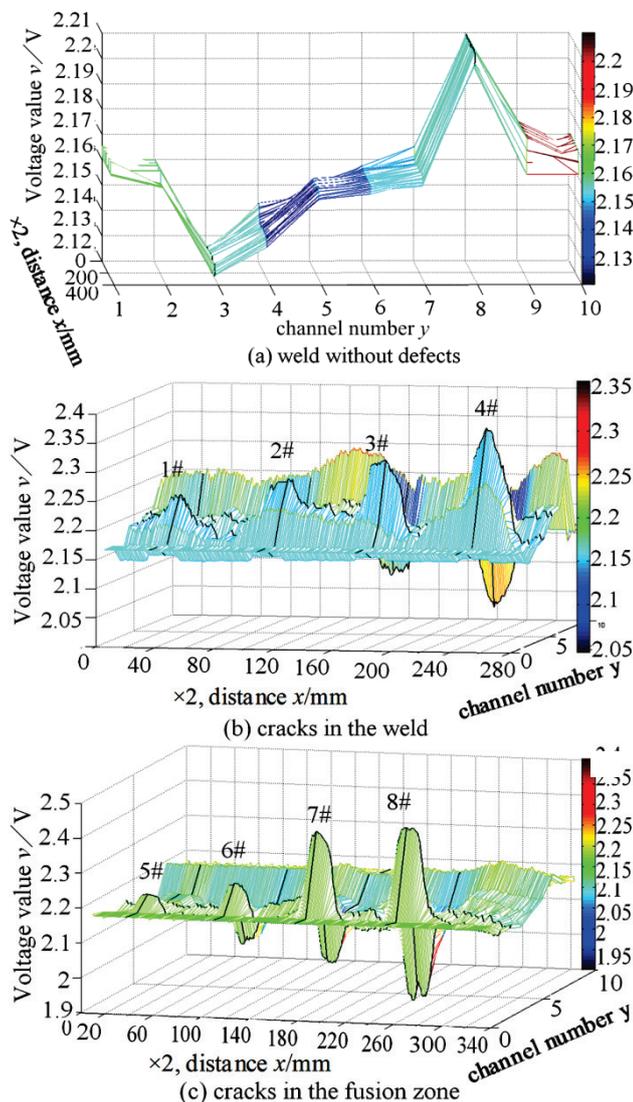


Figure 5 3D maps of MFL signal under the three states

### 3.2 Extraction of the GGCM characteristics

Following the algorithm steps, the three-dimensional distribution of MFL signals were converted into a two-dimensional grayscale diagram. To extract the graphics characteristics accurately by linear gradation transformation, the original grey value was mapped to [0, 255], the MFL data size of the different weld defects corresponded to the grey colour depth, and then the MFL signal data was changed to display the grayscale. A set of grayscale figures under the three kinds of conditions is respectively shown in Fig. 6(a), Fig. 7(a), Fig. 8(a). Using three types of each defect image in each group as a sample, the extracted GGCM of each sample is shown in Tab. 1.

### 3.3 k-means clustering

k-means clustering was first put forward by Mac Queen. Using  $k$  as a parameter, on the basis of minimizing the error function, the data is divided into  $k$  classes in advance, making one category with high similarities, and different categories with lower similarity degrees.

Through k-means analysis of 72 samples of 15 characteristic quantities ( $T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8, T_9, T_{10}, T_{11}, T_{12}, T_{13}, T_{14}, T_{15}$ ) of the above three kinds of weld condition, the results are shown in Tab. 2 and Tab. 3. In Tab. 2 the third column represents the sample belonging to each kind, and the fourth column represents the distance between the sample and the centre of its class.

From Tab. 2, compare the initial sample tag in the clustering results. A crack in the fusion zone (Category 3) is wrongly assigned to cracks in the weld (Category 2), and other samples are separated correctly. And the classification result shows that this method has obtained higher clustering accuracy.

From Tab. 3, the distance between Class 1 and other Classes is bigger than the distance between other Classes. This shows that identifying pipe weld with defects and pipe weld without defects is easier than identifying between pipe weld defects.

Table 1 Fifteen characteristic quantities of GGCM of the three weld kinds

	Pipe weld without defects			Cracks in the weld			Cracks in the fusion zone		
	1	...	24	1	...	24	1	...	24
$T_1$	0.3434	...	0.3649	0.4046	...	0.3915	0.4912	...	0.4315
$T_2$	17.2133	...	15.567	16.4498	...	13.959	13.1748	...	13.1695
$T_3$	24.1111	...	22.4422	58.1781	...	57.4251	29.0129	...	34.6242
$T_4$	141.251	...	116.1882	194.9984	...	177.8501	207.0033	...	178.7745
$T_5$	0.0358	...	0.0388	0.0552	...	0.0453	0.0299	...	0.0291
$T_6$	104.0113	...	104.7897	91.3519	...	104.3261	132.6683	...	122.1863
$T_7$	17.2133	...	15.567	16.4498	...	13.824	13.1291	...	13.3353
$T_8$	63.0031	...	61.8097	62.7493	...	42.6589	66.7509	...	35.7047
$T_9$	14.1166	...	13.6861	14.0184	...	13.6277	14.1524	...	13.7573
$T_{10}$	-494.207	...	-501.881	-206.451	...	-177.954	-548.4041	...	-258.742
$T_{11}$	1.5583	...	1.5346	1.2387	...	1.4812	1.5401	...	1.4851
$T_{12}$	0.6499	...	0.8131	0.5523	...	0.7062	0.6123	...	0.7319
$T_{13}$	1.8166	...	1.7645	1.5524	...	1.7929	1.8074	...	1.8891
$T_{14}$	12492.51	...	12939.88	10071.23	...	12415.24	19903.6389	...	14105.88
$T_{15}$	0.0008	...	0.0016	0.0016	...	0.0022	0.0032	...	0.0006

**Table 2** Cluster members

Sample/No.	Initial sample tag	Category/species	Distance/mm
1	Pipe weld without defects (Category 1)	1	4.825
2	Pipe weld without defects (Category 1)	1	2.281
3	Pipe weld without defects (Category 1)	1	6.592
4	Pipe weld without defects (Category 1)	1	2.719
...	...	...	...
24	Pipe weld without defects (Category 1)	1	2.228
25	Cracks in the weld (Category 2)	2	7.903
26	Cracks in the weld (Category 2)	2	16.212
27	Cracks in the weld (Category 2)	2	5.334
28	Cracks in the weld (Category 2)	2	10.159
...	...	...	...
48	Cracks in the weld (Category 2)	2	13.562
49	Cracks in the fusion zone (Category 3)	2	22.413
50	Cracks in the fusion zone (Category 3)	3	7.517
51	Cracks in the fusion zone (Category 3)	3	13.519
52	Cracks in the fusion zone (Category 3)	3	11.134
...	...	...	...
72	Cracks in the fusion zone (Category 3)	3	10.302

**Table 3** Distance between centers of final clustering

Category	1	2	3
1		45.408	45.421
2	45.408		32.122
3	45.421	32.122	

**3.4 Edge detection for weld defects**

On the basis of the classification of previous steps, edge detection is carried out. As shown in Fig. 6(a), the contrast between the black line with the weld through channel after grayscale linear transformation and the white line with the weld peak channel after grayscale linear transformation was clear. The edge detection is according to the eighth step of the algorithm, with the result shown in Fig. 6(b).

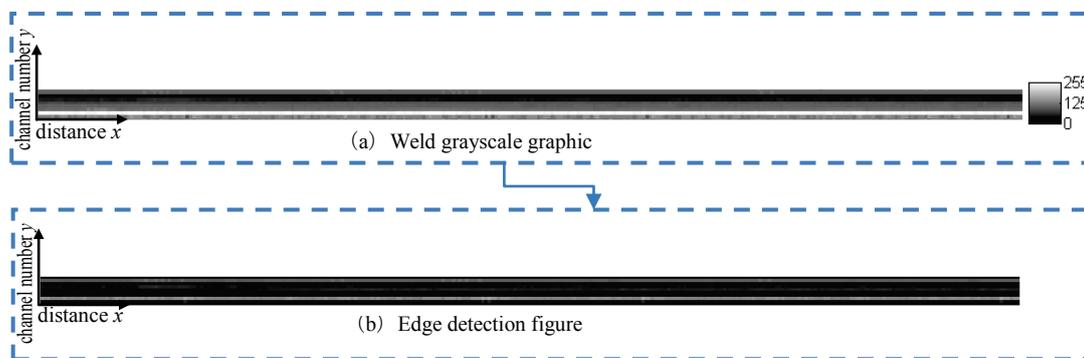
The cracks in different locations of the weld (cracks in the weld, cracks in the fusion zone) are shown in Figs. 7(a) and 8(a). From Figs. 7(a) and 8(a), the differences between the peaks and troughs of the MFL signals increased along with the contrast between the white line with the peak of the cracks after grayscale linear transformation. The information of the weld and the defect can be shown and compared with the MFL signals 3D figure graphics, which is more clearly illustrated by the grayscale graphics.

According to the fifth step of the algorithm, on the basis of Figs. 7(a) and 8(a), the results of histogram equalization are shown in Figs. 7(b) and 8(b). From Figs. 7(b) and 8(b), the overall contrast in the image improved, and the dynamic range of the pixel grey value expanded, which enhanced the overall contrast of the image.

According to the sixth step of the algorithm, binarization threshold was selected by using Otsu's method on the basis of Figs. 7(b) and 8(b), to process the image for binarization, then the results are shown in Figs. 7(c) and 8(c). After binarization processing, the black line (trough) for the cracks in the weld was separated from the background effectively along with the white line (peak) in the fusion zone.)

According to the seventh step of the algorithm, using to remove small objects on the basis of Figs. 7(c) and 8(c), and then the results are shown in Figs. 7(d) and 8(d).

According to the eighth step of the algorithm, by the dilation edge detection of Figs. 7(d) and 8(d), the edge detection results are shown in Figs. 7(e) and 8(e). From Figs. 7(e) and 8(e), the results matched the cracks in the weld and the cracks in the fusion zone. Through the edge detection for weld defects, the results were reliable and had good consistency. It can intuitively judge the distribution and location for the pipe weld cracks, and the experimental results verified the feasibility of the structured algorithm.



**Figure 6** Weld defect-free

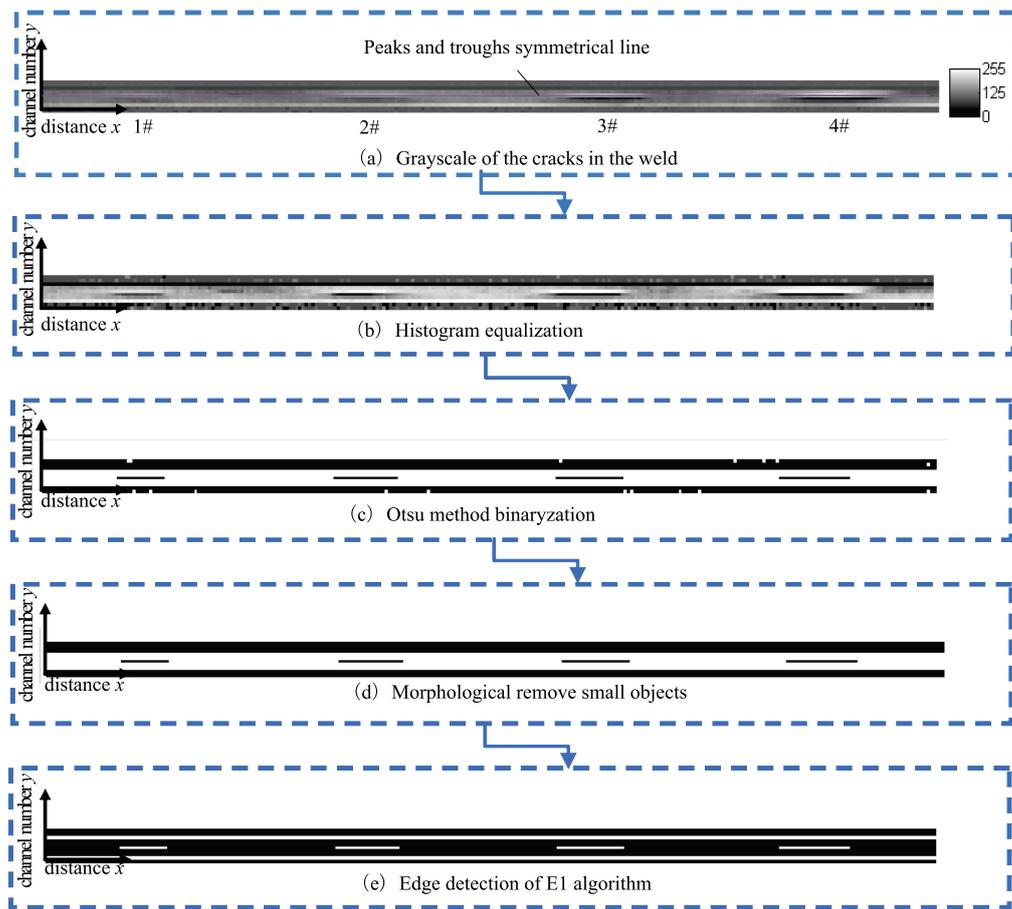


Figure 7 Cracks in the weld

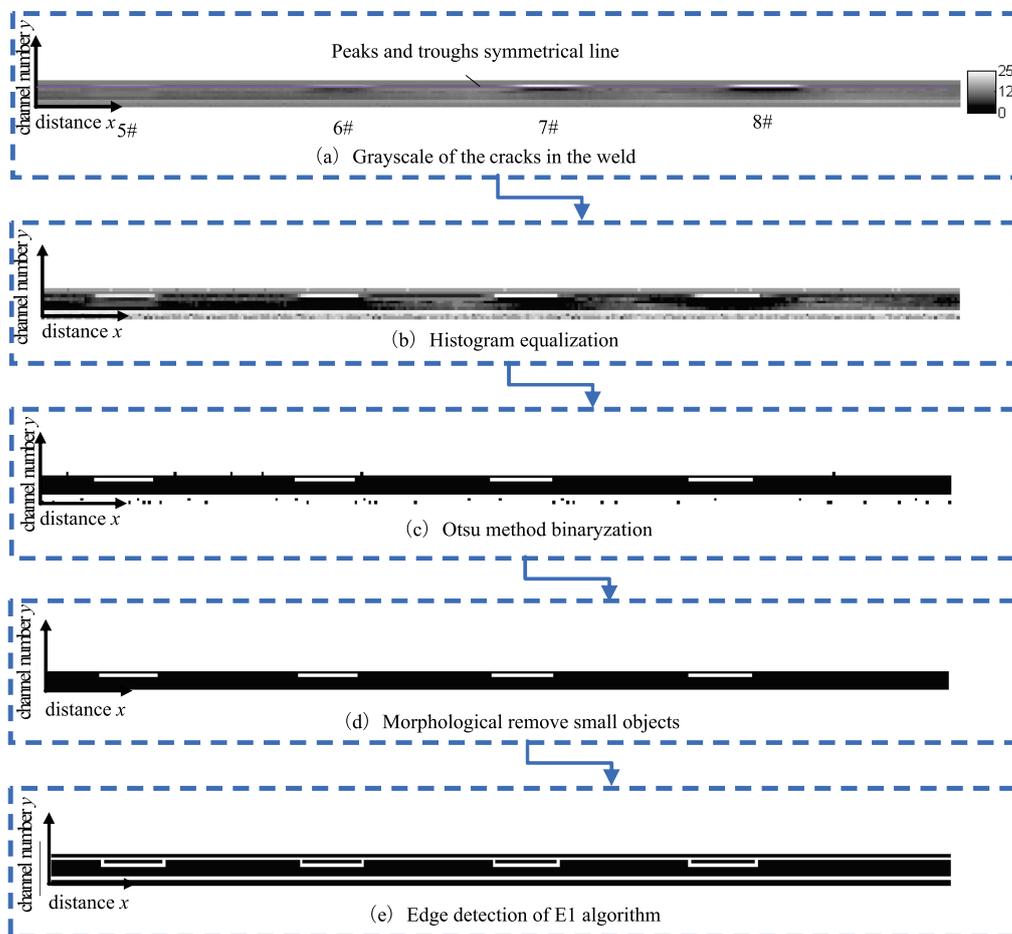


Figure 8 Cracks in the fusion zone

## 4 Conclusion

This paper proposes an image recognition method based on GGCM and cluster analysis and mathematical morphology, which is different from the traditional ideas based on MFL signal processing. By using GGCM and clustering algorithm, this method can identify the pipe weld states effectively for welds without defects, cracks in the pipe weld and cracks in the fusion zone. The results have shown that this method has higher clustering accuracy. On the basis of these classification results, by constructing a mathematical morphology optimization method, the edges of the weld shape contour and the cracks in different areas of the weld (in the weld and in the fusion zone) were extracted. The edges were clear, and had good edge graphics features.

According to the algorithm results, the reliability and validity of approaches proposed have been verified, and break through the applicable limitations of traditional signal processing technology. Later work can explore new ways of pipe weld defect recognition analysis.

## Acknowledgements

This work was sponsored by the National Natural Science Foundation of China (51607035, 11502051) and Heilongjiang Postdoctoral Foundation (LBH-Z16040) and State Administration of Work Safety Science and Technology Project of Key Technologies for Preventing and Controlling Major Accidents in Safe Production (heilongjiang-0003-2017AQ) and Science and Technology Project of China Petroleum and Chemical Industry Association (2017-11-04) and Research start-up fund of Northeast Petroleum University (rc201732) and Petro China Innovation Foundation (Project Number: 2015D-5006-0602). All these are gratefully appreciated.

## 5 References

- [1] Cruz, F. C.; Simas Filho, E. F.; Albuquerque, M. C.; Silva, I. C.; Farias, C. T. Efficient feature selection for neural network based detection of flaws in steel welded joints using ultrasound testing. // *Ultrasonics*. 73, 1(2017), pp. 1-8. <https://doi.org/10.1016/j.ultras.2016.08.017>
- [2] Huggett, D. J.; Dewan, M. W.; Wahab, M. A.; Okeil, A.; Liao, T. W. Phased Array Ultrasonic Testing for Post-Weld and OnLine Detection of Friction Stir Welding Defects. // *Research in Nondestructive Evaluation*. 27, 9(2016), pp. 1-24.
- [3] Boateng, F. T.; Ewert, U.; Kannengiesser, T.; Zscherpel, U.; Griesche, A. Real-time radiography for observation of crack growth during GTA welding. // *Welding in the World*. 60, 9(2016), pp. 931-937. <https://doi.org/10.1007/s40194-016-0351-7>
- [4] Xu, K. S.; Qiu, X. Q.; Jiang, H.; Wei, R.; Zhong, J. Analysis of magnetic memory signal of 20# steel welding defects. // *Hanjie Xuebao*. 37, 3(2016), pp. 13-16.
- [5] Feng, Q. S.; Song, H. C.; Wang, X. L.; Wang, F. X. The Triaxial Magnetic Flux Leakage Signal Analysis of Spiral Weld Roof Topping and Hi-Low Defects. // *Nondestructive testing*. 31, 5(2009), pp. 340-344.
- [6] Kang, Y. H.; Sun, Y. H.; Song, K. The Feasibility Analysis of MFL Method for ERW Tube. // *Journal of Test and Measurement Technology*. 24, 2(2010), pp. 99-104.
- [7] Demir, N. Various building detection methods with the use of image and LIDAR data. // *Tehnički vjesnik*. 21, 2(2014), pp. 341-349.
- [8] Hong, J. G. Gray level-gradient co-occurrence matrix texture analysis method. // *Acta Automatica Sinica*. 10, 1(1984), pp. 22-24.
- [9] Zhu, H. Q. The Segmentation Method of Retinal Blood Vessels Based on Gray Level-Gradient Co-occurrence Matrix. // *Journal of Shanghai Jiaotong University*. 38, 9(2004), pp. 1485-1488.
- [10] Siqueira, F. R. D.; Schwartz, W. R.; Pedrini, H. Multi-scale gray level co-occurrence matrices for texture description. // *Neurocomputing*. 120, 10(2013), pp. 336-345. <https://doi.org/10.1016/j.neucom.2012.09.042>
- [11] Manivannan, K.; Aggarwal, P.; Devabhaktuni, V.; Kumar, A.; Nims, D.; Bhattacharya, P. Particulate matter characterization by gray level co-occurrence matrix based support vector machines. // *Journal of Hazardous Materials*. 223, 2(2012), pp. 94-103. <https://doi.org/10.1016/j.jhazmat.2012.04.056>
- [12] Malge, P. S.; Nadaf, R. S. PCB Defect Detection, Classification and Localization using Mathematical Morphology and Image Processing Tools. // *International Journal of Computer Applications*. 87, 9(2014), pp. 40-45.
- [13] Namdari, F.; Salehi, M. High-Speed Protection Scheme Based on Initial Current Traveling Wave for Transmission Lines Employing Mathematical Morphology. // *IEEE Transactions on Power Delivery*. 32, 1(2017), pp. 246-253. <https://doi.org/10.1109/TPWRD.2016.2571341>
- [14] Bibiloni, P.; González-Hidalgo, M.; Massanet, S. General-Purpose Curvilinear Object Detection with Fuzzy Mathematical Morphology. // *Applied Soft Computing*. 60, 11(2017), pp. 655-669. <https://doi.org/10.1016/j.asoc.2017.07.010>
- [15] Bai, X.; Liu, H. Edge enhanced morphology for infrared image analysis. // *Infrared Physics & Technology*. 80, 1(2017), pp. 44-57. <https://doi.org/10.1016/j.infrared.2016.11.011>

### Authors' addresses

#### *Wei Cui, associate professor*

Northeast Petroleum University,  
Daqing high-tech development zone university street No. 99,  
Daqing City, Heilongjiang Province, China  
E-mail: cuiweivv@126.com

#### *Ke Wang, postgraduate*

Northeast Petroleum University,  
Daqing high-tech development zone university street No. 99,  
Daqing City, Heilongjiang Province, China  
E-mail: 83926003@qq.com

#### *Qiang Zhang, associate professor*

(Corresponding author)  
Northeast Petroleum University,  
Daqing high-tech development zone university street No. 99,  
Daqing City, Heilongjiang Province, China  
E-mail: zqzqzhangq@126.com

#### *Peng Zhang, engineer*

Petrochina Daqing Petrochemical Company,  
Mobile Equipment Department,  
Longfeng district, Daqing City, Heilongjiang Province, China  
E-mail: zpeng2003@126.com