STATISTICAL DETECTION OF DEFECTS IN RADIOGRAPHIC IMAGES OF WELDS

D. Kazantsev[1], G. Salov[1], and V. Pyatkin[1]
[1] Institute of Computational Mathematics and Mathematical Geophysics, Novosibirsk, Russia

Abstract: In this paper we investigate applicability of statistical techniques for defect detection in radiographic images of welds. The defect detection procedure consists in a statistical hypothesis testing using several nonparametric tests. A comparison of rules derived for image thresholding for a given level of false alarm is presented. In this work we consider longitudinal defects such as linear flaws similar to cracks. Numerical experiments with real and test data within MATLAB's abilities are performed.

Introduction: Film radiography is a traditional imaging inspection technique for nondestructive examination of industrial equipment components in order to locate any cavities, inclusions, lack of fusion and so on that may have been formed during the manufacturing or operation process [3,5]. Sophisticated image analysis of digitized films and digital radiographs is a widely studied research field, with much recent approaches using adaptive thresholding and probabilistic methods [1], neural networks [9], texture analysis [11], a wavelet–based multi–resolution image representation [15], a model–based statistical segmentation, mathematical morphology and pattern recognition. We refer to the works [2, 6, 7, and 13] where reviews with a comparative study of the methods can be found. The objective of this work is as follows: given a false alarm level, detect line–like defects in welds in realistic computer time. We try to investigate the possibilities of automatic image processing of weld defects with the help of statistical hypothesis testing using nonparametric statistical tests. We consider several tests which theoretically, for a given level of false alarm, provide us with a threshold resulting in a map of possible defects. The software is implemented in MATLAB and tested on images of welds of austenitic pipes acquired by radiography.

We consider the detection problem as a problem of hypothesis testing, in the most important practical formulation – the absence of a priori information about brightness distribution in the points of object and background. We assume that an image under investigation is inhomogeneous and anisotropic and all observed variables have continuous probability distribution functions. In case the image does not contain an object, we suggest that the observations are statistically independent. Let us define a shape of an investigated object by the form of the window (square, circular, longitudinal, etc.) which moves along the image. For each possible defect location, size and orientation of the window we investigate a problem of defect detection by testing the null hypothesis H_0 : "Absence of a defect" against the alternative H_1 : "Presence of the defect". Rejection or acceptance of H_0 is based on values of some test statistics computed within the sliding window. Hypothesis H_0 is rejected when values of the statistics achieve a critical level (threshold).

Results: For each current pixel we choose longitudinal windows W_1 and W_2 with parameters R_1 and R_2 (Fig. 1). Let $x_1,...,x_N$ be arbitrary pixel values from W_1 and $y_1,...,y_M$ be arbitrary pixels from W_2.
Denote $x = \{x_i\}, i = 1,...,N$; $y = \{y_j\}, j = 1,...,M$. For detecting the object we test the null hypothesis H_0 :"values x_i and y_j are stochastically equal" against the two – sided alternative hypothesis H_1 :" x_i are stochastically larger (or less) than y_j ". For testing, we need appropriate statistics. We are going to test the hypothesis H_0 with the use of statistics which do not depend on the form of the image brightness distribution functions unknown to the observer. Such statistics and the corresponding tests are often called nonparametric.

Sign Test: One of the simplest nonparametric tests is a sign test [12]. Put $M = N$ and assume that the pixels from the object are stochastically brighter (larger) that those from the background (one sided alternative). The Sign Criterion for testing the null hypothesis H_0 is based on the following statistic:

$$\nu = \sum_{i=1}^{N} I\{x_i - y_i > 0\}, \qquad (1)$$

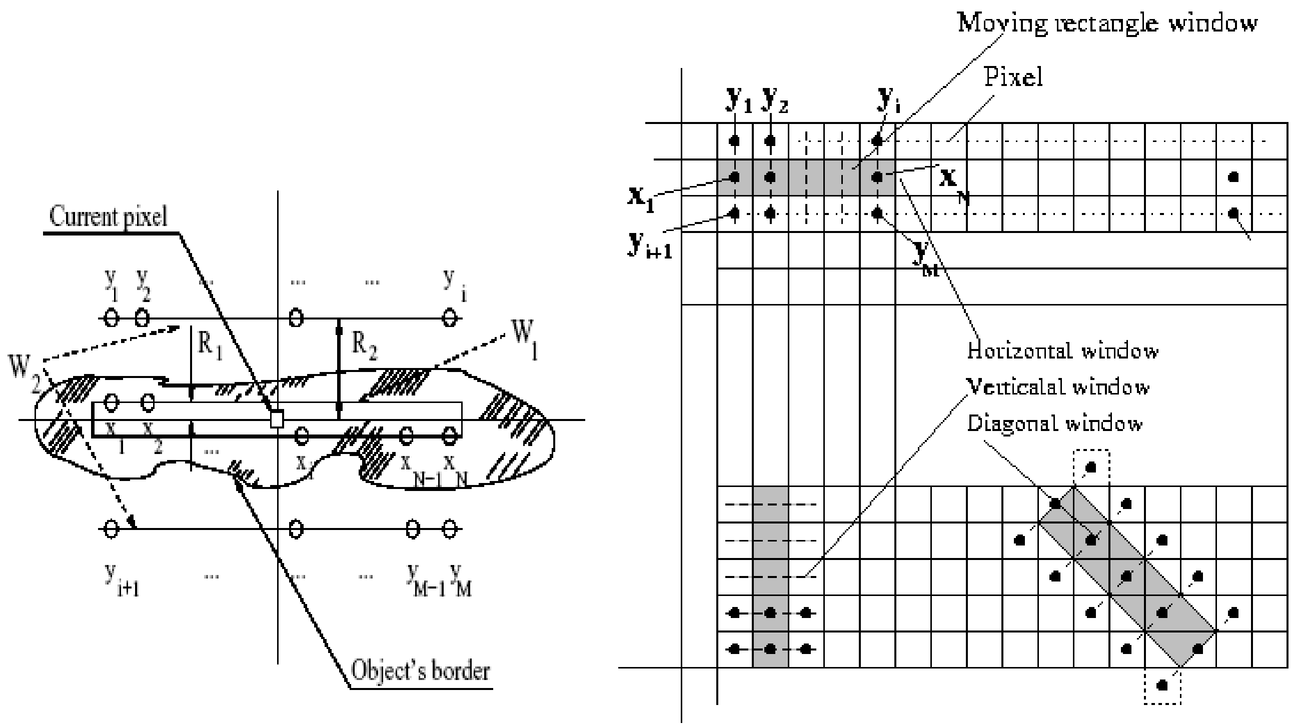

Figure 1: (Left) Longitudinal (horizontal) scheme of scanning. (Right) Image matrix with different orientations of scanning window.

That is the number of positive signs among $x_1 - y_1, \ldots, x_n - y_n$. Here and in what follows, we define *I {A} = 1* if event *A* occurs and *I {A} = 0* if it does not. The sign test rejects the null hypothesis H_0 when $\nu \geq \lambda,$ where the threshold $\lambda = \lambda(\alpha),$ is determined by the acceptable level α and equals the smallest integer such that

$$\sum_{i=\lambda}^{N} \binom{i}{N} 2^{-N} \leq \alpha. \tag{2}$$

Rosenbaum Test: As a second example, consider the Rosenbaum statistics [14]. Denote

$$A_1 = \text{ the number of } x \text{ larger than } y_{\max} = \max\{y_j\}. \tag{3}$$

Given the integer *r,* it can be shown that the probability of event $\{A_1 \geq r\}$ under the null hypothesis H_0 has the form

$$\Pr(A_1 \geq r \mid H_0) = \frac{(N+M-r)!N!}{(M+N)!(N-r)!}. \tag{4}$$

The one-sided Rosenbaum test is as follows. Reject H_0 only if $A_1 \geq \lambda$, where the threshold λ is the smallest integer such that

$$\frac{(N+M-\lambda)!N!}{(M+N)!(N-\lambda)!} \leq \lambda. \tag{5}$$

Let

$A_2 =$ the number of y larger than $x_{\max} = \max\{x_i\}$. (6)

The two – sided Rosenbaum test for H_0 against the two-sided alternative H_1 is as follows. Reject H_0 only if $T_1 \equiv A_1 - A_2 \geq C_1$ or $T_1 \leq C_2$.

Critical levels C_1 and C_2 depend upon the entire number Z of tests. In case of a single test, the probability of false alarm $\Pr(A_1 \geq C_1 \mid H_0) + \Pr(A_2 \geq C_2 \mid H_0)$ is expressed as a ratio

$$\delta = \frac{(N+M-C_1)!N!}{(M+N)!(N-C_1)!} + \frac{(N+M-C_2)!M!}{(M+N)!(M-C_2)!}, \tag{7}$$

where C_1 and C_2 should be chosen so that the fractions in (7) are approximately the same and they are the smallest integers for which the relationship

$$\delta \leq \alpha \tag{8}$$

is hold.

If $M = N$, then $C_1 = C_2$ and calculations are essentially simpler. In this case, given the probability of false alarm

α we reject the hypothesis H_0 when $|T_1| \geq C_1$,

where the threshold C_1 is chosen as the smallest integer such that

$$2\frac{(2N-C_1)!N!}{(2N)!(N-C_1)!} \leq \alpha. \tag{9}$$

Haga Test: The Rosenbaum test can be modified. Let

$B_1 =$ the number of y smaller than $x_{\min}$, $B_2 =$ the number of x smaller than $y_{\min}$.

Then the one – sided Haga test for H_0 against the one – sided alternative H_1 as is follows. Reject H_0 only if $A_1 + B_1 \geq C(\alpha)$. The two - sided Haga test for H_0 is to reject H_0 if

$$T_2 = |A_1 + B_1 - A_2 - B_2| \geq C(\alpha/2). \tag{10}$$

Critical levels $C(\alpha)$ and $C(\alpha/2)$ can be taken from tables [4]

Wilcoxon – Mann – Whitney test: The following nonparametric test for H_0 was originally proposed by Wilcoxon [16], Mann and Whitney [10]. The Mann - Whitney test is based on statistics

$$U^+ = \sum_{i=1}^{N}\sum_{j=1}^{M} I\{x_i - y_j > 0\},$$

$$U^- = \sum_{i=1}^{N}\sum_{j=1}^{M} I\{x_i - y_j < 0\}, \tag{11}$$

($U^+ =$ number of pair $\{x_i, y_j\}$ with $x_i > y_j$). Values U^+ and U^- are integers from 0 up to *NM* and distributed symmetrically near the point *NM/2* when H_0 is true.

The Wilcoxon test is based on statistics

$$W = U^+ + \frac{1}{2}N(N+1). \tag{12}$$

Let the integer $u > 0$ be fixed, and then the probability $P(U^+ = u \mid H_0)(P(U^- = u \mid H_0))$ can be represented in the form

$$P(U^+ = u \mid H_0) = \frac{M!N!}{(M+N)!} \sum_{p(M)} 1, \tag{13}$$

where the summation is performed over all partitions $p(M) = \{m_0, m_1, ..., m_N\}$ of the number M into $N+1$ integer nonnegative addenda's $M = m_0 + m_1 + ... + m_N$
satisfying the equality

$$\sum_{i=1}^{N} i m_i = u. \tag{14}$$

The two – sided Mann – Whitney test for H_0 against the two – sided alternative H_1 is as follows. Reject H_0 only if $U^+ \geq C_U(\alpha/2)$ or $U^- \geq C_U(\alpha/2)$. The one – sided Mann – Whitney test to reject H_0 if $U^+ \geq C_U(\alpha)$. The one – sided Wilcoxon test to reject H_0 if $W \geq C_w(\alpha)$. In this work we use the two - sided Mann – Whitney test
$U = \max\{U^+, U^-\}$.

Discussion: In our preliminary numerical experiments, we used two ways of thresholding:
1) Blind or maximal thresholding, when only maximally possible values for test statistics are used. It means that we visualize only the objects that theoretically correspond to a minimal level of false alarm, i.e. extreme suspects on defect ness.
2) For a given level of false alarm we compute thresholds for different tests with fixed values of such processing parameters as M,N, R_1, R_2 and Z (the entire number of all tests of hypothesis H_0).

The experiments show that empirically chosen thresholds often are very close to the upper limit of the test statistic values. We show the result of detection by the Mann-Whitney test U thresholded by C=6525 with parameters M=N=30, R_1=1, R_2 =3 in Figure 2. Choosing a value of the significance level α , the user should take into account that this choice relates to the expectation of about αZ defects. For instance, if the user intends to escape noisy false defects (but with a risk to miss a true defect), small values of α have to be chosen. We can derive from the experiments that active and convenient interaction with computer and data is needed to tune statistical methods.

For this purpose, we use MATLAB (a product of the Math Works Inc.) as a means of running statistical detection experiments. MATLAB is a powerful platform for high - performance mathematical computation and graphical representation, whose basic element is an N-dimensional matrix. MATLAB provides a wide range of functions for use in data analysis, incorporates a programming language that is similar to many common languages such as FORTRAN, PASCAL and C, and enables a user to communicate with the computer through Graphical User Interface (GUI). The GUI created to accompany the defect detection experiments consists of a system of windows (a main window pictured in Figure 3). A GUI is a useful tool for dealing with thresholding problems and new defect detection approaches training.

Conclusions: The results presented here demonstrate that the methodology of hypothesis testing based on nonparametric statistics can be applied to problems of defect detection with some hope of success. The benefits in detection of sizes of the defects remain somewhat limited, although we do see images with reliably detected defective zones. Even unsupervised thresholding of the map of the investigated test statistics discloses the defects that are difficult to indicate by visual radiographic inspection.

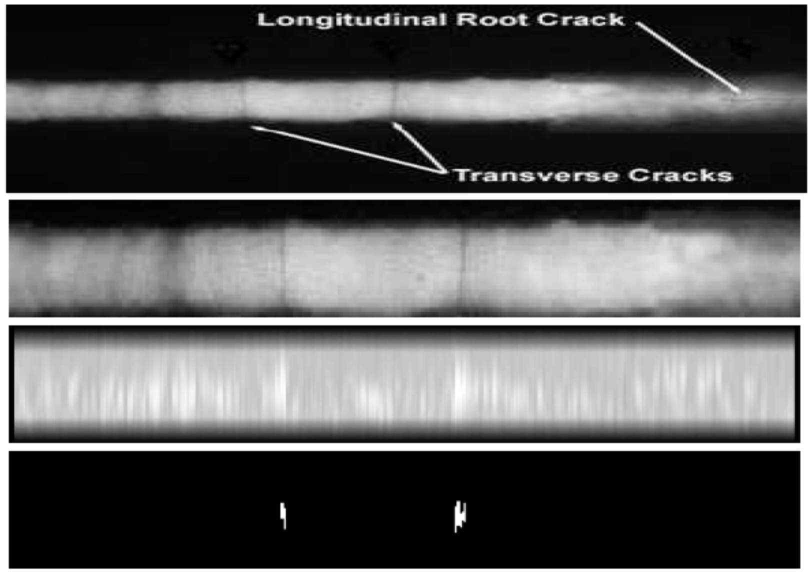

Figure 2: Image of 66 x 440 sizes with two transverse cracks; map of the Mann–Whitney test U statistic values thresholded by C = 6525, when maximal test value is C = 6700, for $M = N$ = 30, $R_{\min} = 1; R_2 = 3$.

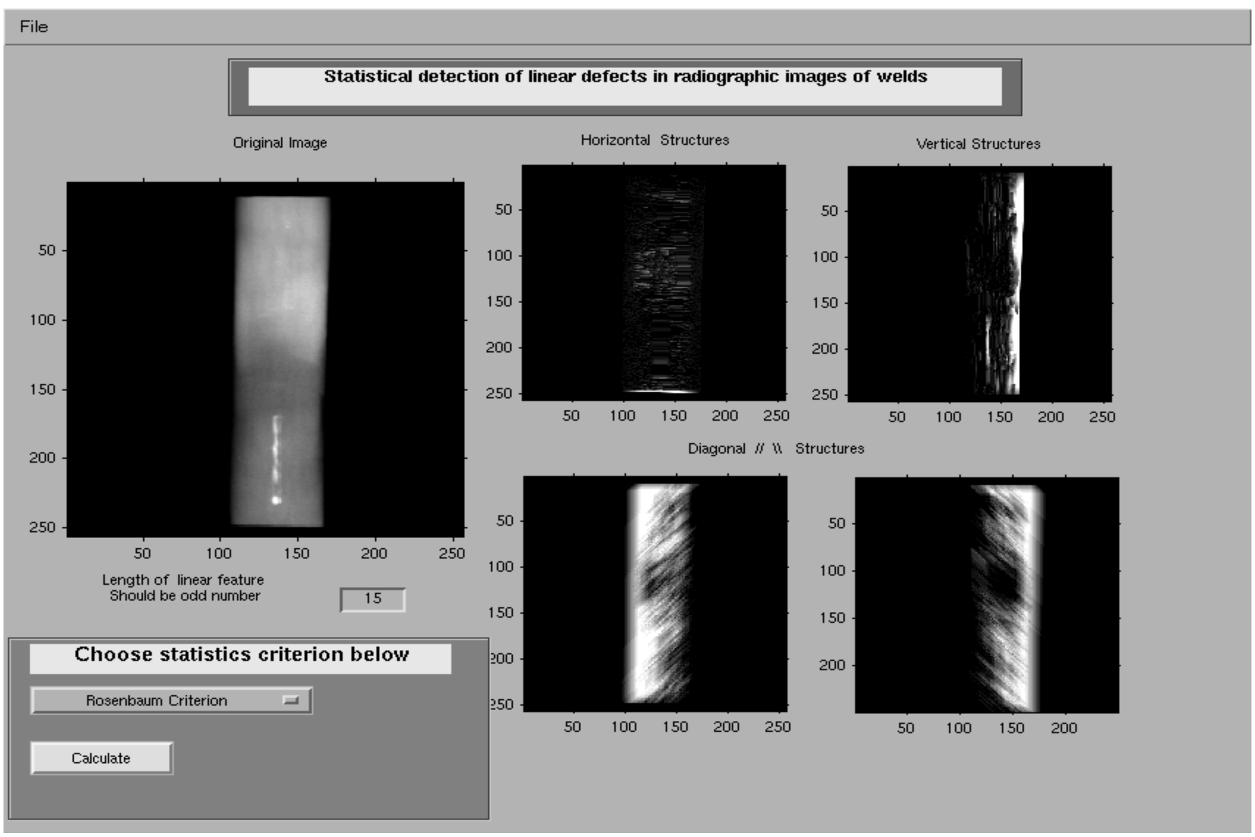

Figure 3: Main window of graphical user interface.

References:

[1] O. Alekseychuk, U. Zscherpel, Probabilistic crack detection in digitized radiographs of weld inspection, Proc. of Int. Symp.on Comp. Tomography amd Image Processing for Industrial Radiology, Berlin, Germany, 2003, pp.211-222.

[2] H. Boerner, H. Strecker, Automated X–Ray Inspection of Aluminium Castings, IEEE Trans. on Patt. Anal. Mach. Intell. 10 (1) (January 1988) pp. 79-91.

[3] L.E. Bryant, P. McIntire, R.C. McMaster (Eds), Nondestructive Testing Handbook, Second Edition, Volume 3, Radiography and Radiation Testing, ASNT, ISBN 0-931403-00-6, 1985.

[4] T. Haga, A two-sample rank test on location, Ann. Inst. Math. Stat., 11 (1959/1960), pp. 211-219.

[5] R. Halmshaw , Proceedings of the 14th World Conference on NDT, New Delhi, India, 1996, pp 27-32.

[6] R. Halmshaw, N.J. Ridyard, A review of digital radiological methods, British Journal of NDT 32 (1) (1990) ,pp. 17-26.

[7] V. Kaftandjian, Y.M. Zhu, G. Peix, D. Babot, Automatic recognition of defects inside aluminium ingots by X-ray imaging, Insight 38 (9) (September 1996), pp. 618-625.

[8] I.G. Kazantsev, I. Lemahieu, G.I. Salov and R. Denys, Statistical detection of defects in radiographic images in nondestructive testing, Signal Processing 82 (2002), pp. 791-801.

[9] S.W. Lawson, G.A. Parker, Intelligent segmenation of industrial radiographs using neural networks, Proceedings of SPIE, vol. 2347, 1994, pp. 245-255.

[10] H.B. Mann, D.R. Whitney, On a test of whether one or two random variables is stochastically larger than other, Ann. Math. Stat. 18 (1947), pp. 50-60.

[11] D. Mery, M.A. Berti, Automatic detection of welding defects using texture features, Proc. of Int. Symp.on Comp. Tomography amd Image Processing for Industrial Radiology, Berlin, Germany, 2003, pp. 281-290.

[12] H.R. Neave, P.L. Worsington,Distribution-free Tests, Unwin, London, 1988.

[13] R.M. Palenichka, U. Zscherpel, Detection of Flaws in Radiographic Images by Hypothesis Generation and Testing Approach, Proc. Int. Workshop Advances in Signal Processing for Nondestructive Evaluation of Materials, X. Maldague (Ed.), ASNT, Topics on NDE(TONE), 1998, pp. 235-243.

[14] S. Rosenbaum, Tables for a nonparametric test of location, Ann. Math. Stat. 25 (1954), pp. 146-150.

[15] R.K. Strickland, H.I. Hahn, Wavelet transform methods for objects detection and recovery, IEEE Transactions on Image Processing 6 (5) (1997), pp. 724-735.

[16] F. Wilcoxon, Individual comparisons by ranking methods, Biometrics Bull. 1 (1945), pp. 80-83.