CrossMark

The Application of Rough Sets Theory to Design of Weld Defect Classifiers

Tomasz Chady[1] (iD) · **Ryszard Sikora[1]** · **Leszek Misztal[1]** · **Barbara Grochowalska[1]** · **Bogdan Grzywacz[1]** · **Michał Szydłowski[2]** · **Paweł Waszczuk[1]** · **Mariusz Szwagiel[3]**

Abstract The innovative method for weld defect classification based on rough set theory is presented in this study. The classification rules have been generated by processing of data base composed of 640 radiographic images referring to certain welding process in aircraft industry. The obtained accuracy of defect identification (from 88% up to 100%, depending on class of defect and choice of classifier) can be evaluated as at least competitive or even better one comparing to results referring to other type of frequently "exploited" classifiers, those mentioned in attached overview section. The identification of weld defects is the final operation which is premised by complicated "chain" of consecutive operations transforming primary radiographs to the form enabling calculation of conditional attributes. That is why brief description of process of transformation of primary radiographs to the forms which are suitable for attributes calculation is included in the paper.

1 Introduction

The radiographic NDT methods of welds inspection are commonly used in order to evaluate the quality of welds. In period since 1990 till now we can observe considerable efforts aimed at automatic analysis of radiographs. The current paper concentrates on the idea and possible applications of promising method for classification of welds defects based on rough set theory [1–4]. The usefulness of proposed method for computerized classification of weld flaws was confirmed by introductory experiments and comparative data presented in [5,6]. The method allows to determine the reducts, i.e. minimal set of attributes which preserve ability of distinguishing the classes of weld defects. Thus, the redundant attributes can be removed by means of formalized procedures. The discussion of other problems encountered during realization of consecutive operations preparing input image data to the form which allows coming to decision on weld quality and class of imperfection is marginally included in scope of current paper. Let us note that hundreds of publications yielded during last 20 years have been devoted to these problems. The automatic classification of weld defects is based on input data generated by chain of previously done operations, i.e. weld images acquisition and preliminary processing, segmentation, feature extraction, detection of flaws. Thus, even faultless classifying algorithms do not guarantee perfect final results for "poor" input data caused by errors and omissions contributed by former transformations of primary image. The standard EN ISO 6520-1 distinguishes between five types of welding defects, i.e. cracks, slag inclusions, porosity, lack of fusion, and continuous undercuts. The levels of acceptance of welds containing those imperfections are determined in standard EN ISO 5817 and the other ones.

Depending on sizes and other parameters referring to geometric and texture properties of single defect or cluster of defects the classifying algorithm should recognize them as acceptable or not. If necessary, the weld flaws should be assigned to classes defined by standards mentioned above.

✉ Tomasz Chady
tchady@zut.edu.pl

1 Faculty of Electrical Engineering, West Pomeranian University of Technology, Szczecin, Poland

2 Faculty of Mechanical Engineering and Mechatronics, West Pomeranian University of Technology, Szczecin, Poland

3 Pratt & Whitney S.A., Rzeszow, Poland

In order to design classifying algorithms one has to choose the set of parameters (descriptors, conditional attributes) as basis for decision making process and select the type of classifier processing these parameters. The computational complexity can be important agent influencing the choice of classifying algorithm as well. If we take into account quite substantial number of known type of classifiers and procedures determining classifier parameters as well as data drawn from published reports, where up to about 60 different attributes describing weld flaws were processed, then the following conclusion seems to be obvious: design of classification algorithm occurs as task characterized by large number of "degrees of freedom". That is why researchers proposed variety of approaches. It could be argued that almost all existing data mining methods and algorithms for classification have been used by scientific teams struggling with maximizing of classification accuracy. The still existing importance of problem discussed in current paper is proved by tens, if not hundreds, papers and reports dealing with welds flaws detection, segmentation and classification published every year since the early nineties of the last century to the present days.

2 The Brief Overview of Methods Used for Weld Defects Classification

The numerous approaches have been used for design of weld defects classifiers. It seems, that most of them were designed by means of methods belonging to class of statistical and/or AI ones. Basing on the arbitrary and rough review of attempts to classification given in this section one is able to recognize what are the typical algorithms which are useful for automatic classification of weld defects. Furthermore, the overview provides with brief information on accuracies of various classifiers as well as number and "essence"of input parameters used in classification procedures. The short information on less or more complicated methods of determining of relevant and redundant parameters can be found as well. Nevertheless, the advanced presentation of contemporary "state of art" ought to be the subject of comprehensive, dedicated publication which does not fall in framework of current paper.

In recent paper [7] the Bayesian Networks (BN) based on probability theory has been used as classifying tool. Choosing the invariant geometric attributes like in [8], i.e. compactness, elongation index, rectangularity, symmetry, deviation index to the largest inscribed circle, Euclidian lengthening, etc. and learning the BN with set composed of several hundreds of welds samples the promising results have been obtained. For "fifty/fifty" splitting of data into learning and testing sets the defects, like cracks (CR), lack of penetration (LP), porosities (PO), slug inclusions (SI), were classified correctly with accuracy 90–97%.

The expectation-maximization (EM) statistical method of classification of weld defects is presented in [9]. The data clustering EM algorithm is an iterative way finding maximum likelihood or maximum *a'posteriori* estimates of parameters of statistical models, where model depends on latent variables. The results yielded by EM algorithm were compared to results obtained by artificial neural network (ANN) classifier with input represented by 4 principal components created on the basis of several "elementary" features and learned by error backpropagation technique. The presented experimental results showed that performance of ANN classifier (accuracy about 97%) is a little bit better comparing to results of EM method (accuracy about 92%).

In [8] problem of relationship between elementary geometric parameters of weld defects and classes of defects has been considered. This analysis inclined authors to introducing of "hybrid" parameters SGD (Shape Geometric Descriptor) and GFD (Generic Fourier Descriptor). Furthermore, the combined descriptor $f_{(CFD,FGD)}$ has been proposed in order to discriminate better the problematic defects. The use of hybrid descriptors is especially advised when CBIR (Content-Base Image Retrieval) technique of classification is used. This technique consists in comparison of currently examined weld defect to correctly classified "patterns" of defects collected in data base. It is worth to be mentioned that procedure of creation of "hybrid" parameters on the basis of primary parameters is one among commonly used advanced tools aimed at deletion of inconsistencies in decision tables.

The very interesting study of statistical approach to weld defects classification on the basis of texture features is presented in [10]. The following two groups of widely used texture features are taken into consideration:
(1) features based on the co-occurrence matrix, which gives a parameterized information of how often one grey value will appear in a specified spatial relationship to another grey value on the image (angular second moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measures of correlation 1, information measures of correlation 2, maximal correlation coefficient),
(2) features based on 2D Gabor functions.

The selection of features for classification from the initial set of 64 based on Gabor filtering and 148 based on co-occurrence matrix were done by means of the SFS (Sequential Forward Selection) method and analysis of ROC (Receiver Operation Characteristics) curves. The paper states that best texture features drawn from co-occurrence matrix were the mean of the difference entropy and the mean of the difference variance, for a distance of $d=3$, while the best Gabor's features were those obtained for scale $p=6$, and orientations: \,—, /. The classification was

executed by means statistical classifiers, i.e. polynomial, Mahalanobis and nearest neighbor. The sensitivities (defined as TP/(TP+FN), where TP—number correctly classified defects, FN—number of flaws recognized as non-defect) yielded by above classifiers were about 91% except those for nearest neighbor procedure, where sensitivities were less than 80%. This evaluation has been obtained after examination of almost 1400 samples including 200 with defects.

The AI techniques can be concerned as most popular tools used to classification of weld defects. The classic expert system with weld defects classification rules drawn from sophisticated decision tree is described in [11]. The implemented knowledge base was gathered from specialists, textbooks and international standards. The expert system can identify the 11 faults in gas pipelines welded by shielded metal arc welding. The classification of flaws is carried out on the basis of features concerning shape of defect (rectangularity factor, quotient of perimeter and area), orientation (size in X direction, size in Y direction), location (inside weld, centre of weld, edge of weld, base metal). The dimensions of classified weld defect are compared with requirements of standards (API, ASME, DIN, BS, AWS, ABS, JIS) which allow to formulate final decision on weld acceptance or rejection. The discussed paper does not contain the numbers characterizing accuracy of classification. Nevertheless, it can be found concluding statement that accuracy does not differ from this yielded by top experts.

The other rule-based expert system for detection of flaws (however without procedures for classifying of them) has been described in [12]. This system seems to be one among pioneering, where attributes of transversal gray level line profiles and "technology" of curve fitting have been used to find the weld defective areas.

The fuzzy expert systems presented in [13,14] can be treated as classic examples of classifier based on fuzzy-logic (FL) approach. The so-called WM (Wang, Mendel) machine learning method [15] was used to generation of fuzzy rules from more than 100 examples. The choice of number of partitions for input and output data was supported by means of genetic algorithm (GA). The other fuzzy rules based classification algorithm for attributes selected by means of principal component analysis (PCA) was presented in [6]. Depending on class the obtained MSE (Mean Square Errors) were in range 0.02–0.05. The best results were achieved for cracks, which was the most separated group in case of PCA analysis. This algorithm was included in software of Intelligent System for Radiographs Analysis (ISAR). On the beginning authors of ISAR system took into consideration almost 60 weld defect attributes: 22 representing geometrical and textural properties, 6 representing brightness and 32 representing central, normalized and invariant Hu's shape moments [16–18]. The "basic" classifier included in ISAR system was that in form of Multi-Layer Perceptron Neural Network

(MLPNN) with 2 hidden layers containing up to 16 neurons [19]. The database for learning and testing contained about 1500 samples of weld defects representing standard classes of weld imperfections. The comparative examinations of results given by MLPNN classifier and classifiers based on k-means clustering algorithms proved that accuracy was much better in case of applying of ANN (depending on class and number of samples for ANN learning and testing the proper classification coefficients were in the acceptable range 0.6–0.7).

Artificial neural networks are commonly chosen tools for classification of patterns based on learning from examples. In researches reported in [20] the set of 43 descriptors corresponding to texture measurements (angular 2nd moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance var., data from co-occurrence matrix) and geometrical features (relative position to the weld bead, the aspect ratio, the length/area ratio, the area/bounding rectangle ratio, the roundness, the rectangle ratio, the Heywood diameter and the relative angle to the weld bead) were determined for each segmented object and put as classifier input. The classifiers were trained to classify each of the objects into one of the defect classes or indicate lack of defect. Three-fold cross validation was carried out for evaluation of experimental results. The paper contains comparison of results yielded by ANN-classifier to results obtained by use of SVM (Support Vector Machine) and k-NN (k-Nearest Neighbourhood) classifiers. The experiments were carried out for 411 segments (objects) corresponding to worm holes (85 cases), porosity (94 cases), linear slag inclusion (42 cases), gas pores (13 cases), lack of fusion (57 cases), crack (26 cases) and non-defects, i.e. false positives (94 cases). To reject redundant data the Sequential Backward Selection (SBS) method has been applied. The obtained accuracies were about 87% for both ANN and SVM classifiers and identical input data represented by 43 descriptors. After diminishing the number of descriptors to 7 the ANN classifier preserved almost the same accuracy as in case of 43 descriptors but accuracy of SVM was considerably lower. For the k-NN classifier the results were significantly inferior.

The interesting results concerning nonlinear ANN classifiers are presented in [21,22]. The ANN was composed of two layers (intermediate and output one). The number of neurons in intermediate layer was optimized, i.e. the net performance and its training errors were taken into account to find the optimal number of neurons. Like previously, the problem of relevance of input data has been examined. The criterion for choice of relevant attributes was based on observations of changes in the network responses when a feature (attribute) used was substituted by its average value. The larger the difference between these net responses, the larger attribute relevance. The dimension of input vector initially represented with four components, has been diminished gradually to 3,

2 and 1 components in function of the features selected by relevance criteria. The paper shows how to apply the NN network to determining the first principal nonlinear discrimination component and consecutive ones. The presented idea can be treated as generalization of PCA (Principal Component Analysis) which in fact is widely used techniques for the size reduction of multivariable data set. The criterion of relevance together with nonlinear classifiers proved that only four of the six initial features were relevant for the classification of defects. The obtained accuracy was about 99–100% for training set containing 125 samples described by 4 conditional attributes each. Thus, the considered paper supports thesis that "quality" of the conditional attributes is more important than their "quantity". By the way: problem of relations between parameters quantity and quality is discussed even in "historical" papers [23,24] where authors convince that no more than 10 meaningful features is sufficient for accurate weld defect classification.

The problems associated with design of NN classifiers of weld flaws are deeply discussed in [25]. The samples for 140 non-defects, 126 slag inclusions, 87 porosities, 8 transversal and 14 longitudinal cracks were examined. The initial set of attributes was composed of 12 members represented area, centroid (*x* and *y* coordinates), major axis, minor axis, eccentricity, orientation, Euler number, equivalent diameter, solidity, extent and position. The samples with defect and non-defect ones were classified by outstanding experts. The ANN classifiers were learned by backpropagation method using Widrow–Hoff algorithm for multiple-layer networks and non-linear differentiable neuron transfer functions. To eliminate the redundant attributes the PCA transformation of input data has been executed. In order to overcome problem of ANN overfitting the procedures of Bayesian regularization and bootstrap approaches have been implemented. Furthermore, the classic MSE (Mean Square Error) criterion evaluating accuracy of ANN classifier was supplemented with additional term representing the mean of the sum of squares of the network weights and biases. The paper contains comparative results of classifications for Bayesian and bootstrap regularization as well as classic and supplemented MSE criteria of ANN performance. Authors of [25] conclude that the proposed technique is capable of achieving good results. The ANN (composed of 11 input neurons and 20 neurons in hidden layer) optimized by means of supplemented MSE criterion and PCA transformation of input data has occurred to be the best classifier. Depending on class of defect the accuracy of classification fell in the range from 72% (slug inclusion) up to 96% (cracks). The non-defect samples were classified correctly with accuracy 92% and obtained mean accuracy was about 80%.

There are numerous reports dealing with classifiers in form of feedforward multilayer neural networks [26,27] where supervised learning of ANN was executed and error

backpropagation learning was often used. The classifiers based on networks of ART (Adaptive Resonance Theory) type are presented in [28]. The networks of this type are examples of classifiers which can be learned without supervision, i.e. algorithm of learning generates patterns of defects on the basis of "distances" between vectors of attributes describing weld flaws. It leads finally to determining of centers of clusters and so-called largest radiuses of similarity recommended for each class of defects. Analysis of so-called demerit factors gave possibilities of creation of several pattern curves (2–9) for each class of defects. The examination of classifier under consideration was carried out with use of attributes drawn from processed gray level curves of weld samples representing 100 transversal profiles for each concerned classes (lack of discontinuity, porosity, longitudinal crack, slug inclusion, lack of fusion, lack of penetration, undercutting).

In [29] were compared accuracies of NN classifiers to those based on pruned and non-pruned decision trees optimized by measure of entropy. The following types of NN were used: a backpropagation net, a RBF (Radial Basis Function) Net, a Fuzzy ARTMAP Net and LVQ (Learning Vector Quantization) Net. The 36 features were initially taken into account. The best accuracies calculated for testing set were obtained for RBF and backpropagation net - both in the range 94–95% of correct classifications. The error rates of other type classifiers were in range 9–12%.

At the end of this section it should be noted, that authors of papers often neglect presentation of complete information about structure of algorithms and their parameters. Moreover, the described examinations were not done for identical databases. That is why comparisons of final accuracies of methods presented in accessible publications should be done very carefully.

3 The Classification of Defects Based on Rough Sets Theory

The processing of weld defect image yields set of parameterized attributes (see former section) which can be useful for defect classification purposes. The data referring to images of defects can be gathered in decision table which can be treated as representation of information system (SI). The developed in the last twenty years of previous century theory of rough sets [1–4] can be considered both as extension of classic sets theory and tool enabling successful design of classifiers.

The discretization of conditional attributes usually diminishes the number of samples (rows) in decision table because parameters of certain subsets of elements ("classes of abstraction") belong to the same ranges created for values of attributes by discretization. The modified decision table obtained due to discretization can contain redundant

attributes as well. The so-called "*B-indiscernibility relation*" can be used as precisely defined tool enabling the rejection of redundant attributes. Let $SI = (U, A)$ be information system, and $B \subseteq A$, where U is considered set of samples ("universe"), A, B is set (subset) of attributes respectively. The relation:

$$IND(B) = \{(x, y) \quad \in U \times U : \forall \alpha \in B, \alpha(x) = \alpha(y)\} \tag{1}$$

is called the „*B-indiscernibility relation*". Thus, the elements x, y from U which fulfill (1) are indescirnible because all values of attributes α belonging to subset B and representing those elements are identical. If subset B induces classes of abstraction which are identical to those generated by A then attributes of A in classification procedures can be substituted with reduced number of attributes, i.e. by attributes belonging to B.

The rough sets theory introduces the lower and upper approximations of set. On this basis one can evaluate the level of „belief" concerning the membership of given object to set D (representing decision, concept, idea, *type of weld defect*, etc.). The lower approximation:

$$B_{IND(B)}D = \{x \in U : [x]IND(B) \subseteq D\} \tag{2}$$

represents elements of universe U belonging to classes of abstraction induced by attributes of subset B where all members of these classes fulfill conditions of membership to set D. The upper approximation of D is defined as:

$$B^{IND(B)}D = \{x \in U : [x]IND(B) \cap D \neq \phi\} \tag{3}$$

and represents all classes of abstraction induced by B which contain any elements or even single element belonging to D. The boundary region:

$$BN_{(B)}(D) = B^{IND(B)}D - B_{IND(B)}D \tag{4}$$

is composed of elements belonging to these classes of abstraction where values of attributes from set B do not allow to decide if given element belong to D or not. For example, if we assume that decision D denotes "unacceptable crack in weld" then exact rules of classification of this defect have to be based on lower approximation of D. For introductory selection of elements which potentially can be defective ("cracked") one can use knowledge defining the upper approximation of D. The elements (welds) of boundary region can be cracked or not. For boundary region elements, the task of overcoming of classification inconsistencies has to be solved (through the introduction of new conditional attributes, changing of parameters for attributes discretization, introduction of "artificial" parameters by creation of

functions defined on previously used parameters, etc.). Of course, depending on particular requirements, one can agree for certain level of faulty decisions using other solutions, like widely known "simple voting" approach, probabilistic rule induction [30]. The other exemplary algorithms processing inconsistent Decision Tables can be found in [30–33]. The presented LEM2 (Learning from Examples Module, version 2) rule induction algorithm uses rough set theory to handle inconsistent data set. LEM2 algorithm induces a set of certain rules from the lower approximation and a set of possible rules from the upper approximation. The procedure for inducing the rules is the same in both cases - see [32]. The algorithm imitates the classical greedy scheme yielding a local covering of each examples from the given approximation using a minimal set of rules [33]. The preliminary discretization of attributes is not necessary when algorithm called MOD-LEM is used [32]. The MODLEM processes these attributes during rule induction, when elementary conditions of a rule are created. The MODLEM algorithm can operate according to its Entropy Version or Laplace Version. MODLEM uses rough set theory to handle inconsistent examples and determinates a single local covering for each approximation of the concept [32]. The search space for MODLEM is bigger comparing to search space for LEM2, because LEM2 creates rules basing on already discretized attributes. That is why rules obtained by MODLEM are "simpler and stronger". The algorithm called EXPLORE extracts all decision rules which satisfy requirements, referring to strength, level of discrimination, length of rules and syntax conditions of rules. It can be adapted to processing of inconsistent examples by using rough set approach or by tuning a proper value of the discrimination level. Induction of rules is executed as result of exploring the rule space with simultaneous imposing the restrictions referring to above mentioned requirements. Procedure of exploration of the rule space is repeated for each concept to be decided. Each concept can be a class of examples or one of its rough approximations in case of inconsistent examples. The "kernel" of algorithms was built using breadth-first exploration of rule space, beginning from one-condition rules. Exploration of given branch is stopped if the requirements are satisfied or stopping condition is fulfilled (impossibility of fulfilling the requirements is attained [33]).

The existence of mentioned and several other algorithms inclines to using of rough sets approach to classification of weld defects. Thus, definitions (2), (3), (4) can be taken into consideration during organizing of process of inducing of classification rules, because depending on results of rough selection (determination of upper approximation) the farther classification procedures can be activated. The existing programs, like "Rough Set Exploration System" (http://logic.mimuw.edu.pl/~rses/get.htm), can be helpful tools for classification experiments.

Table 1 The exemplary information system

$A \rightarrow$ $U \downarrow$	$a1$	$a2$	$a3$	D
$u1$	0	0	0	$D1$
$u2$	0	1	0	$D2$
$u3$	0	0	1	$D3$
$u4$	1	1	0	$D2$
$u5$	1	1	1	$D4$
$u6$	1	0	1	$D5$

By rejection of redundant attributes we obtain "reduct", i.e. minimal subset of attributes that enables the same classification of elements of the universe U as the whole set of attributes. Usually one obtains set of reducts. Every reduct (set of attributes) contains certain subset of attributes which is identical for all reducts. It is called "core". If sets $IND(B)$ $=IND(A)$, then $CORE(B) = \cap RED(A)$, where $RED(A)$—set of all reducts classifying elements of U identically like with use of all attributes of A. The parameters of discretization influence substantially the sizes of sets (2),(3),(4) as well as forms and number of reducts. In general, the problem of determining of reducts for big sizes of Decision Tabels is considered as non-deterministic polinomial-time hard (NP-hard) one. That is why genetic algorithms or other heuristics are used for determining of reducts. Fortunately, the conclusions and advices given in many papers (see previous section) contain statements that weld defect classification should be done with use of not to big number of relevant parameters.

Let us illustrate the "philosophy" of idea of creating of classification rules by elementary example. For Information System defined by Table 1 we can create Discernibility Matrix (Table 2) and Functions, or other equivalent representations (without redundant data like in symmetric Discernibility Matrix) which are more convenient for computer processing. The Discernibility Matrix shows these attributes which obtains different values for elements belonging to couple defined by raw and column labels. Thus, one obtains information which attributes are necessary in order to distinguish objects of U, or classes of abstractions u_i, if u_i represents class of objects fulfilling the given indiscernibility relation. Let us note, that SI in Table 1 is deterministic one, because it does not contain identical sets of values of conditional attributes which generate different decisions. The lack of such property makes SI non-deterministic.

For example, object (or class) $u2$ differs from $u5$ because of values of attributes $a1$, $a3$. The discernibility function for $u2$ is:

$$f2 = a2 \wedge (a2 \vee a3) \wedge a1 \wedge (a1 \vee a3) \wedge (a1 \vee a2 \vee a3)$$
$$= ... = (a1 \vee a2) \wedge (a2 \vee a3)$$

This function shows attributes which differ $u2$ from other objects of U. For $u4$ is:

$$f4 = (a1 \vee a2) \wedge a1 \wedge (a1 \vee a2 \vee a3) \wedge a3 \wedge (a2 \vee a3)$$
$$= ... = (a1 \vee a3) \wedge (a2 \vee a3)$$

Hence the decision rule classifying objects to decision class $D = D2$ is:

$$IF\ (a1 = 0\ AND\ a2 = 1)\ OR\ (a2 = 1\ AND\ a3 = 0)$$
$$OR\ (a1 = 1\ AND\ a3 = 0)\ THEN\ D = D2$$

Let us note that *CORE* of SI in Table 1 is composed of attributes {$a1, a2, a3$} because of known property, that *CORE* contains all these attributes which exist in "cells" of discernibility matrix as single attribute. This means, that number of attributes can not be reduced (all attributes are necessary) if the same classification like in Table 1 has to be kept.

The procedure of data processing for weld defect classification is shown in Fig. 1. The algorithms LERS (Learning from Examples based on Rough Sets) similarly to other approaches based on "learning from examples" which process Decision Tables of big sizes (numerous objects, attributes and ranges of attributes) yield the satisfying classification rules after execution of iterative process where parameters for discretization of attributes, choice of subsets of accessible attributes, choice of objects to learning and testing sets are changed many times. The welds defects used for learning of classifier have to be assigned to standard classes of defects by top experts. The research process of classification with the usage of rough sets theory consists of the six steps, those represented in Fig. 1 as well as by following stages:

Table 2 The discernibility matrix

	$u1$	$u2$	$u3$	$u4$	$u5$	$u6$
$u1$	–	$a2$	$a3$	$a1, a2$	$a1, a2, a3$	$a1, a3$
$u2$	$a2$	–	$a2, a3$	$a1$	$a1, a3$	$a1, a2, a3$
$u3$	$a3$	$a2, a3$	–	$a1, a2, a3$	$a1, a2$	$a1$
$u4$	$a1, a2$	$a1$	$a1, a2, a3$	–	$a3$	$a2, a3$
$u5$	$a1, a2, a3$	$a1, a3$	$a1, a2$	$a3$	–	$a2$
$u6$	$a1, a3$	$a1, a2, a3$	$a1$	$a2, a3$	$a2$	–
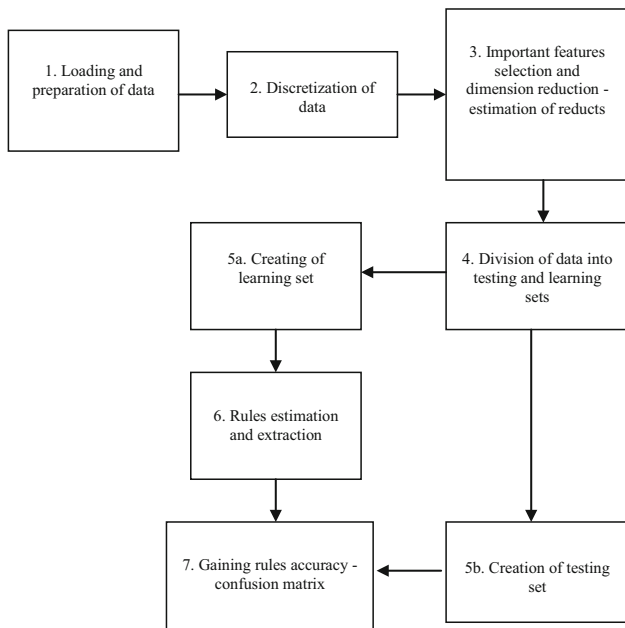
Fig. 1 Procedure for weld imperfections identification

Fig. 2 The cropped weld radiograph

Fig. 3 *B*-Spline reconstructed weld image

Fig. 4 Absolute difference image

- data retrieval from calculated parameters on recognized weld imperfections,
- discretization of attributes,
- computation of attribute importance and decision aimed at dimension reduction,
- dividing of information table into learning and testing sets,
- process of extracting rule sets,
- computation of confusion matrix and other qualitative coefficients.

The commonly used measures of classification accuracy can be applied to results generated by procedures based on rough sets approach. Thus, one can calculate "basic" evaluating parameters, like true positive *TP*, false positive *FP,* true negative *TN,* false negative *FN,* positive $Pos = TP + FP$, negative $Neg = TN + FN$, true $True = TP + TN$, false $False = FP + FN$. Depending on needs one can also determine the "derived" parameters: sensitivity (recall) $Se = TP/(TP + FN)$, specificity $Sp = TN/(TN + FP)$, precision $P = TP/(TP + FP)$, measure $F_b = [(b^2 + 1)TP]/[(b^2 + 1)T]P + FP + b^2FN$, accuracy $Acc = (TP + TN)/(Pos + Neg)$, the balanced accuracy $BAcc = c^1(TP(Pos)^1 + TN(Neg)^{-1}))$ and some others.

4 The Classification of Weld Defects on the Basis of Real Data

The main problem discussed in current paper concerns the possibilities of using of algorithms drawn from rough sets
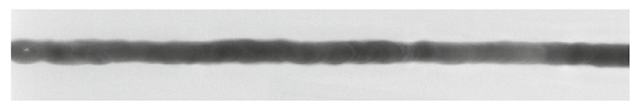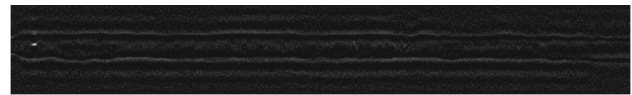
theory to classification of weld defects. Nevertheless, assignment of defects to class defined by standards can be treated as the "terminal" stage which is executed after several operations processing primary radiographs image. Thus, some brief information on preparing data to design of weld defect classifiers is necessary. On the beginning all radiographs were normalized. This operation fits the distribution of illumination of processed image to distribution represented by known parameters (mean, standard deviation). Then algorithm of weld detection and its shape extraction is applied. There are following operations included to this algorithm: low pass filtering of image (reduction of noise and texture details), local tresholding of filtered image with use of sliding window in accordance to Nibalac's method, removing of morphological noise by closing operation, filtering of obtained binary image on the basis of pixel area criteria (removing of small objects with acceptable sizes in relation to defects sizes accepted by standard requirements), hole filling (creation of weld mask), calculation and extraction of rectangular bounding box with a certain margins from the detected weld. The exemplary cropped and normalized image of the radiograph is shown in Fig. 2. The feature extraction is the next set of operations. The cropped radiograph image is "passed" through median filter (removing of "paper and salt" type of disturbances). Next, the weld is divided into *N*-pixel wide windows. For each window an average profile is computed. Then averaged windowed profiles are fitted by using the *B*-Spline functions (reconstruction of weld profile). The reconstructed image is shown in Fig. 3. Afterwards the calculation of absolute difference between the reconstructed and original image is done (Fig. 4). Finally, the simple thresholding (threshold level is determined on the basis of image mean and STD values) generates the binary images which can be treated as final result of stage under consideration. During the next stage the final filtering by morphological closing reduces the unwanted objects in the image. Next the area based particle

eliminates one-pixel artefacts obtained due to thresholding process. To erase artifacts outside the image the masking technique is applied. Finally, the objects still existing in the image are labeled and "exported" for further processing by means of identification procedures—see Fig. 5. The exemplary image of weld containing two types of defects (crack, pores) is shown in Fig. 6. Result of processing of this image according to above mentioned procedure is shown in Fig. 7.
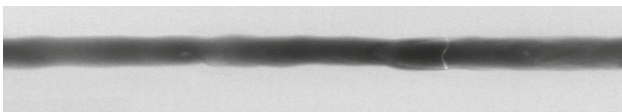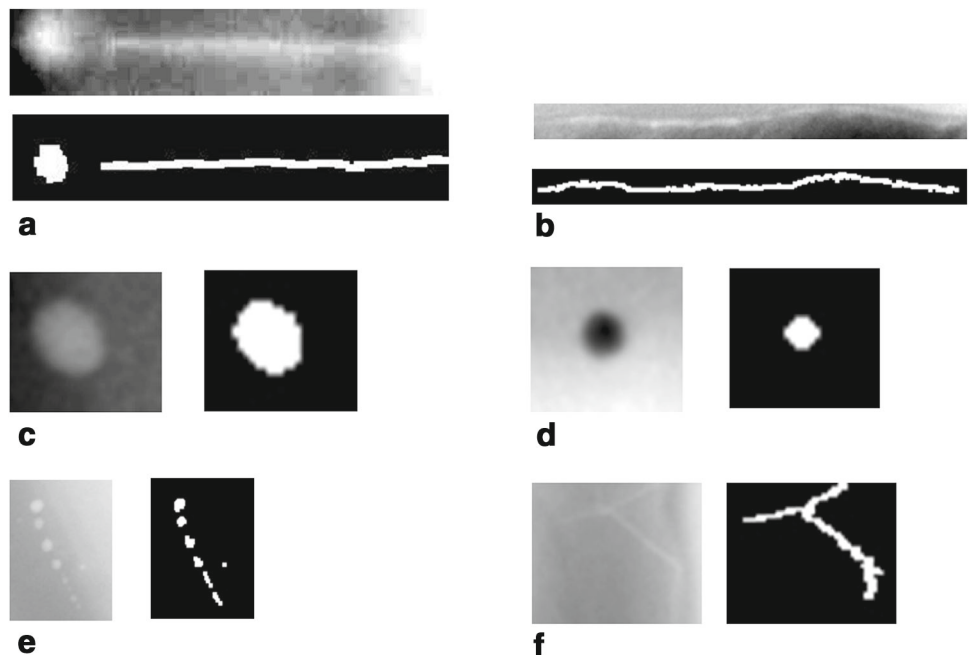

Fig. 5 Detected weld flaws

Fig. 6 Weld containing two types of flaws (crack and pores)

Fig. 7 Detected weld flaws (crack and pores)

The database used later in machine learning process was created on the basis of 640 original radiographs of the thin-walled welded joints of aircraft structures. All the radiographs were achieved from digital radiography system. The flaws have been marked manually by an expert with use of dedicated software application. For each defect its binary mask was created to facilitate computation of the flaw's geometrical attributes. In certain number of cases generating of binary masks was supported by manual correction. The exemplary images of the flaw and its mask for each defect type are presented in Fig. 8. Finally, the images of the flaws cut from the original radiograph along with their masks were saved and the full information about each defect for every radiograph was gathered in resulting *.xml* file.

The statistics of flaw's imperfections in database representing certain process of welding in aircraft industry is presented in Table 3. One can observe strongly differing numbers of examples in consecutive classes of imperfec-

Table 4 The statistics of reduced database dedicated for classification rules induction and extraction

No. of flaws	No. of cracks (class 1)	No. of porosities (class 2)	No. of undercuts (class 5)
152	39	109	4

Fig. 8 Examples of flaws and their binary masks: **a** lack of fusion, **b** undercut, **c** porosity, **d** tungstem inclusion, **e** group of pores, **f** crack

Table 3 The statistics of processed data

No. of flaws	No. of cracks (class 1)	No. of porosities (class 2)	No. of inclusions (class 3)	No. of lacks of fusion (class 4)	No. of undercuts (class 5)
499	39	453	1	1	4

Table 5 The exemplary "piece" of decision table (examples beneath belong to class 1 or 2)

Original image	set1_Sample-08-70kV-15mA-11s-IQI_cut.tif	set1_Sample-08-70kV-15mA-11s-IQI_cut.tif	set1_Sample-05-60kV-12mA-8s-IQI_cut.tif	set1_Sample-09-60kV-12mA-8s-IQI-B_cut.tif	set1_Sample-09-60kV-12mA-8s-IQI-B_cut.tif	set1_Sample-10-60kV-12mA-8s-IQI_cut.tif
Flaw image	922885249.tiff	922885250.tiff	497749372.tiff	101391212.tiff	101391213.tiff	813105053.tiff
Flaw mask	922885249_mask.tiff	922885250_mask.tiff	497749372_mask.tiff	101391212_mask.tiff	101391213_mask.tiff	813105053_mask.tiff
IXX	91822	6121.22	42262.2	26966.1	69371	23391.6
IXY	1919.15	−32.2616	14874.2	1979.26	−738.208	4851.41
IYY	92135.5	8643.32	206588	243124	309543	56590.2
IXXX	−47793.8	178.085	59883.2	−10146.2	−9010.3	31305.7
IXXY	38183.6	−1465.14	2333.74	−3493.57	2948.41	3200.48
IXYY	11783.7	134.45	−202642	−17668.4	−50973.8	−31946.5
IYYY	30850.3	−6376.75	−316580	−38112.9	24384.5	−77736.1
NXX	0.0973887	0.108979	0.0437366	0.0315845	0.0425399	0.0624535
NXY	0.0020355	−0.000574367	0.0153931	0.00231824	−0.000452686	0.0129528
NYY	0.0977212	0.153881	0.213796	0.284764	0.189819	0.151091
NXXX	−0.00162676	0.000205947	0.00197661	−0.00039095	−0.000154619	0.00337866
NXXY	0.00129966	−0.00169438	0.000077032	−0.000134614	0.000050595	0.000345411
NXYY	0.000401083	0.000155485	−0.006688750	−0.000680796	−0.000874721	−0.00344782
NYYY	0.00105005	−0.007737443	−0.0104496	−0.00146856	0.000418442	−0.0083966
H1	0.19511	0.262859	0.257532	0.316348	0.232358	0.213545
H2	0.000016684	0.00201751	0.0298679	0.0641212	0.0216918	0.00852771
H3	0.000016125	0.000005318	0.000599964	0.000003861	0.00000617	0.000277144
H4	0.0000007023	0.000082374	0.000129794	0.000003719	0.00000128	0.000064715
H5	0	0.000000002	0.000000036	0	0	0.000000005
H6	−0.000000022	0.000003695	0.000017530	0.000000376	−0.000000123	0.000005764
H7	0	0	0.000000006	0	0	−0.0000000007
Class of flaw	1	2	1	1	1	1

tions which can be result of applying of restrictive, repeatable procedures during welding. For purpose of design of classifier, the table has to be changed to balance differences between numbers of different type of flaws which is necessary for creation of identification algorithm. The number of porosities has been reduced to 109 and classes referring to inclusion and lack of fusion have been eliminated, because farther researches does not make sense for single examples in these classes (compare the contents in Tables 3, 4). The mentioned above introductory processing of defect images yielded binary in grayscale forms representing defects (see Figs. 5, 7, 8). That is why 21 attributes characterizing shape of collected flaws have been taken into account. The family of attributes was initially composed of the seven moments of inertia I, the seven normalized moments of inertia N and moments of Hu (H) of orders $1 \div 7$, expressed in form of functions of normalized moments of inertia (compare to set of attributes in Table 5).

The researches were aimed at obtaining of acceptable classifiers based on possible small number of attributes. The specificity and scrupulosity of welding process under consideration make that set of forms of defects contains "relatively similar" members belonging to given class of defects. This allowed to diminish substantially the number of relevant attributes up to $2 \div 3$ while 21 attributes were initially taken into account. Furthermore, due to mentioned above rigors the acceptable accuracies of classification were obtained even for checking of values of single attribute. These extremely simple classifiers based on single attribute yielded decision after checking the value I_{xx} or H_1, where I_{xx} - moment of inertia (shift invariant moment about the center of mass of defect image)

$$I_{xx} = \sum xx - \frac{(\sum xx)^2}{A}, \tag{5}$$

H_1-invariant to shift, scale, rotation first order moment of Hu:

$$H_1 = N_{xx} + N_{yy} = \frac{I_{xx}}{A^2} + \frac{I_{yy}}{A^2} \quad . \tag{6}$$

The A in (6) denotes the area of defect representation (number of pixels), N_{xx}, N_{yy} represent normalized moments of inertia, x, y are the coordinates of pixels of defect in relation to axis x or y respectively. According to procedure in Fig. 1 the data has been randomly divided into learning data set (50% of examples) and testing set (50% of examples). The exemplary "piece of input data" covering six examples for extraction of classifying rules is shown in Table 5.

The accuracy of classification obtained with rules generated by means of RSES program (http://logic.mimuw.edu.pl/~rses/get.htm) is described by figures gathered in the confusion matrixes shown in Tables 6, 7. The figures without brackets refer to results of applying of classification rules to

Table 6 The results of classification (confusion matrix for reduct based on I_{xx})

Predicted → Actual ↓	Class 1	Class 2	Class 5	Accuracy [%]	Coverage [%]
Class 1	17	0	0	100%	94.4%
	(38)	(0)	(0)	(100%)	(97,4%)
Class 2	0	50	0	100%	90.9%
	(51)	(384)	(2)	(87,9%)	(96,5%)
Class 5	0	0	0	0%	0%
	(0)	(0)	(1)	(100%)	(25%)
Total accuracy: 100% , (88,9%)					
Total coverage: 88.2% , (95,4%)					

Table 7 The confusion matrix for classification by means of reduct based on H_1

Predicted → Actual ↓	Class 1	Class 2	Class 5	Accuracy [%]	Coverage [%]
Class 1	7	0	0	100%	36.8%
	(27)	(0)	(0)	(100%)	(69,2%)
Class 2	0	50	0	100%	89.3%
	(48)	(358)	(2)	(87,7%)	(90,1%)
Class 5	0	0	1	100%	100%
	(0)	(0)	(4)	(100%)	(100%)
Total accuracy: 100%, (88,6%)					
Total coverage: 76.3%, (88%)					

reduced population of examples, these containing 152 members (see Table 4). The numbers in brackets represent results of applying of the same rules (i.e. drawn from 152 examples) to classification of the entire population of examples (i.e. for 499 examples).

5 Conclusions

Let us note, that results for class 5 have been obtained for processing of data referring to 4 examples. Thus, they have to be treated in category of „curiosity", especially in case where all or almost all examples were included to learning set (see figures without brackets for class 5 in Table 6). The comparison of qualities of results obtained for classes 1 and 2 (accuracy, coverage) to those obtained by use of other methods (see rough overview in Sect. 2) inclines to conclusion, that approach based on rough sets theory is competitive in relation to other methods. The specificity of processing of unbalanced data covering in fact 2 classes of defects does not empowers to draw the general conclusions. Nevertheless, one ought to take into account, that presented good results distinguishing 3 classes of weld imperfections were obtained with use of extremely "low-size" reducts. This inclines to supposition that classifiers of weld defects "adjusted" to specific and rigorous welding process can be based on small number of attributes. Thus, the obtained initial results have to be treated as promising. The final stage of radiograph processing, i.e. classification of weld defects by use of small number of con-

ditional attributes, cannot be treated as very complicated task. Nevertheless, one should keep in mind the complicated way of processing of radiographs to the forms enabling calculation of conditional attributes, like these given by (5) and (6). Presently Authors of current paper try to supplement the data base with more examples referring to imperfection classes 3,4,5. The application of described methodology to processing of big size, balanced data base covering all classes of weld imperfections should lead us successfully to comprehensive evaluation of concepts associated with use of rough sets theory to weld defect classification and in fine to design and implementation of automatic system classifying defects for chosen operations of welding in aircraft industry.

References

1. Pawlak, Z.: Rough Sets—Theoretical Aspects of Reasoning About Data. Kluwer Academic Publishers, Boston (1991)
2. Pawlak, Z.: Some Issues on Rough Sets. Springer, Berlin (2005)
3. Inugushi, M.: Generalizations of Rough Sets and Rule Extraction. Springer, Berlin (2005)
4. Stepaniuk, J.: Rough—Granular Computing in Knowledge Discovery and Data Mining. Springer, Berlin (2008)
5. Sikora, R., Baniukiewicz, P., Chady, T., Psuj, G., Grzywacz, B., Misztal, M.: Artificial neural networks and fuzzy logic in nondestructive evaluation. Stud. Appl. Electromagn. Mech. **38**, 137–151 (2014). doi:10.3233/978-1-61499-354-4-137
6. Sikora, R., Baniukiewicz, P., Chady, T., Grzywacz, B., Lopato, P., Misztal, M., Napierała, L., Piekarczyk, B., Pietrusewicz, T., Psuj, G.: The software application and classification algorithms for welds radiograms analysis. In: AIP Conference Proceedings **1511**, 1104–1111 (2013). doi:10.1063/1.4789166. Accessed 25 Jan
7. Goumeidane, A., Bouzaieni, A., Nacerdoline, N., Tobbone, S.: Bayesian networks-based defect classes discrimination in weld radiographic images. Lect. Notes Comput. Sci. **9257**(2), 554–565 (2015)
8. Nacereddine, N., Ziou, D., Hammam, L.: Fusion-based shape descriptor for weld defect radiographic retrieval. Int. J. Adv. Manuf. Technol. **68**(99–12), 2815–2832 (2013)
9. Nacereddine, N., Tridi, M., Hamami, L., Ziou, D.: Statistical tools for weld defect evaluation in radiographic testing. ECNDT (2006) http://www.ndt.net/article/ecndt2006/doc/Mo.2.4.4.pdf
10. Mery, D., Berti, M.: Automatic detection of welding defects using texture features. Insight **45–10**, 776–781 (2003)
11. Shafeek, H., Gadelmawla, E., Abdel-Shafy, A., Elewa, I.: Automatic inspection of gas pipeline welding defects using an expert vision system. NDT&E Int. **37**(4), 301–307 (2004)
12. Liao, T., Li, Y.: An automated radiographic NDT system for weld inspection: part II-flaw detection. NDT&E Int. **31**(3), 183–192 (1998)
13. Liao, T.: Classification of welding flaw types with fuzzy expert systems. Exp. Syst. Appl. **25**, 101–111 (2003)
14. Liao, T.: Fuzzy reasoning based automatic inspection of radiographic welds: weld recognition. J. Intell. Manuf. **15**(1), 69–85 (2004)
15. Wang, L., Mendel, J.: Generating fuzzy rules by learning from examples. IEEE Trans. Syst. Man Cybern. **22**(6), 1414–1427 (1992)
16. Fluser, J., Suk, T., Zitova, B.: Moments and Moment Invariants in Pattern Recognition. Wiley, New York (2009)
17. Teh, C., Chin, R.: On image analysis by the methods of moments. IEEE Trans. Pattern Anal. Mach. Intell. **10**, 496–513 (1988)
18. Hu, M.: Visual pattern recognition by moment invariants. IEEE Trans. Inf. Theory **8**, 179–187 (1962)
19. Sikora, R., Baniukiewicz, P., Chady, T., Lopato, P., Misztal, M., Napierała, L., Pietrusewicz, T., Psuj, G.: Intelligent system for radiogram analysis for welds' defects inspection. In: 18th World Conference on Nondestructive Testing, Durban, South Africa (2012). http://www.ndt.net/article/wcndt2012/papers/609_wcndtfinal00608.pdf. 16–20 Apr 2012
20. Valavanis, I., Kosmopoulos, D.: Multiclass defect detection and classification in weld radiographic images using geometric and texture features. Exp. Syst. Appl. **37**(12), 7606–7614 (2010)
21. Silva, R., Caloba, L., Siqueiraa, M., Rebelloa, J.: Pattern recognition of weld defects detected by radiographic test. NDT&E Int. **37**, 461–470 (2004)
22. Silva, R., Siqueira, M., Caloba, L., Sagrilo, L., Rebello, J.: Evaluation of the relevant characteristic parameters of welding defects and probability of correct classification using linear classifiers. Insight **44**(10), 616–622 (2002)
23. Kato, Y., Okumura, T., Matsui, S., Itoga, K., Harada, T., Sugimoto, K., Michiba, K., Iuchi, S., Kawano, S.: Development of an automatic weld defect identification system for radiographic testing. Weld. Word **30**(7/8), 182–188 (1994)
24. Aoki, K., Suga, Y.: Application of artificial neural network to discrimination of defect type automatic radiographic testing of welds. ISI Int. **39**(10), 1081–1087 (1999)
25. Vilar, R., Zapata, J., Ruiz, R.: An automatic system of classification of weld defects in radiographic images. NDT&E Int. **42**, 467–476 (2009)
26. Silva, R., Mery, D.: State-of-the-art of weld seam radiographic testing: part II Pattern recogtion. NDT Net (2007). http://www.ndt.net/article/v12n09/mery2.pdf. Accessed 1 Sept 2007
27. Padua, G., Silva, R., Mery, D., Rebello, J., Luiz, P., Coloba, L.: Detection and classification of weld discontinuities in radiographic images, part I: supervised learning. Mater. Eval. **65**, 1139–1145 (2007)
28. Padua, G., Silva, R., Mery, D., Rebello, J., Luiz, P., Coloba, L.: Detection and classification of weld discontinuities in radiographic images, part II: unsupervised learning. Mater. Eval., 1230–1236 (2007)
29. Perner, P., Zscherpel, U., Jacobsen, C.: A comparison between neural networks and decision trees based on data from industrial radiographic testing. Pattern Recognit. Lett. **22**, 47–54 (2001)
30. Grzymala-Busse, J., Yao, Y.: Probabilistic rule induction with the LERS data mining system. Int. J. Intell. Syst. **26**, 518–539 (2011)

31. Grzymala-Busse, J.: LERS—A system for learning from examples based on rough sets. In: Slowinski, R. (ed.) Intelligent Decision Support Handbook of Application and Advances of the Rough Sets Theory, pp. 3–18. Kluwer Academic Publishers, Boston (1992)

32. Grzymala-Busse, J., Stefanowski, J.: Three discretization methods for rule induction. Int. J. Intell. Syst. **16**, 29–38 (2001)

33. Stefanowski, J., Vanderpooten, D.: Induction of decision rules in classification and discovery oriented perspectives. Int. J. Intell. Syst. **16**, 13–27 (2001)