



**HAL**  
open science

## Improvement of X-ray castings inspection reliability by using Dempster-Shafer data fusion theory

Ahmad Osman, Valerie Kaftandjian, Ulf Hassler

► **To cite this version:**

Ahmad Osman, Valerie Kaftandjian, Ulf Hassler. Improvement of X-ray castings inspection reliability by using Dempster-Shafer data fusion theory. *Pattern Recognition Letters*, 2011, 32, pp.168-180. hal-00878528

**HAL Id: hal-00878528**

**<https://hal.science/hal-00878528>**

Submitted on 13 Nov 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 1 Improvement of X-ray castings inspection reliability by using Dempster-Shafer data fusion 2 theory

3  
4

5 Ahmad Osman<sup>1,2</sup>, Valérie Kaftandjian<sup>2\*\*</sup>, Ulf Hassler<sup>1</sup>

6

7<sup>1</sup> Fraunhofer Development Center X-ray Technologies, A Cooperative Department of IZFP Saarbrücken and IIS  
8Erlangen, Dr-Mack-Str. 81, 90762 Fürth, Germany.

9<sup>2</sup> National Institute of Applied Sciences INSA- Lyon, Non Destructive Testing using Ionising Radiations laboratory  
10CNDRI, Bat. St. Exupery, 25 Avenue Capelle, 69621 Villeurbanne, France.

11

## 12Abstract:

13

14The aim of this work is to improve the classification of defects in X-ray inspection by developing a  
15new method based on Dempster-Shafer data fusion theory where measured features on the detected  
16objects are considered as information sources. From the histogram of features values on a learning  
17database of manually classified objects, an automatic procedure is proposed to define a set of mass  
18functions for each feature. The spatial repartition of features is divided into regions of confidence  
19with corresponding mass functions. A smooth transition between regions is ensured by using fuzzy  
20membership functions. The whole process is carried out without any expert intervention. Validation  
21takes place on a testing database. Data fusion leads to a significant improvement of classification  
22performances with respect to the actual system.

23

24Key words: X-ray imaging, castings inspection, true defects, false alarms, features extraction, data  
25fusion, confidence levels, mass value, Dempster-Shafer theory.

26

## 271. Introduction:

28

29X-ray inspection is increasingly used as a tool for non destructive testing of industrial parts, such as  
30aluminium castings in the automotive sector. In this field, new materials or processes are developed  
31for cost and weight optimization. As those materials are operated near their load limit, it is neces-  
32sary to check each produced part thoroughly. One main difficulty in X-ray castings inspection is the  
33detection of false alarms (or false defects), especially if very small and low contrasted defects have  
34to be detected. Therefore, efforts have to be made to reduce the rejection rate of good parts without  
35risking to miss true defects. Once the X-ray images are processed for defect detection, the problem  
36can be reduced to a classification problem, in order to discriminate the true defects (TD) from the  
37false defects (FD).

38

---

\* Corresponding author: Tel. +33 4 7243 8061. Email: [valerie.kaftandjian@insa-lyon.fr](mailto:valerie.kaftandjian@insa-lyon.fr) (V. Kaftandjian).

39Several approaches have been proposed for the classification of true defects and false defects, using  
40X-ray inspection.

41Filtering techniques adapted to specific defect shapes such as bidirectional template matching tech-  
42niques were used by (Newman and Jain, 1994) for finding gross casting defects. The method is at-  
43tractive but limited to specific defect types and time consuming.

44  
45Neural networks were used by (Lawson and Parker, 1994) for the automated detection of discon-  
46tinuities in X-ray images of welds. A big advantage of this approach is to learn the neural net on a  
47limited number of defect samples. In another work (Bonser and Lawson, 1998) a defect enhance-  
48ment technique is introduced using a laws filter matched for round defects and Kirsch oriented filter  
49matched for longitudinal defects was reported. The authors concluded that the neural net is more  
50sensitive to fine defects but gives a high false alarm rate, whereas the filtering approach is more re-  
51liable and faster but less sensitive to small defects. Additional investigations of neural networks are  
52described in (Dobrazanski et al., 2005) for flaws in aluminium castings. The optimum network type  
53was selected among multilayer perceptron (MLP), linear network, radial base functions network  
54(RBF), generalized regression neural network (GRNN) and probabilistic neural network (PNN).  
55The best results were obtained employing the PNN and MLP networks with performances of 90 %  
56and 81 % respectively on the test vectors, and 85 % and 94.5 % for the validation vectors. Perform-  
57ance is here given by the classification correctness of different defect types, but the false alarms rate  
58is not mentioned.

59  
60In (Hernandez et al., 2004) the extracted features of each potential flaw were analysed and assigned  
61to one of the following classes: 'defect' or 'regular structure' by using a Neuro-Fuzzy method  
62named ANFIS (Jang, 1993) which is used for constructing and training a fuzzy inference model that  
63best classifies new data. The advantage of Neuro-Fuzzy systems is the combination of both proper-  
64ties: non linear learning based on numerical data and handling uncertainties in data. However the  
65fuzzy model identification is very sensitive to which attributes are presented and class distribution  
66in the training data. The developed system was used in casting and welds defects detection. The best  
67performance was obtained for aluminium castings with 95% of true defect classification (57 out of  
6860 instances of defect class) and less than 1% (199 non defects classified as defects out of 22876  
69available) of false alarms rate. However the number of true defects on which the method is valid-  
70ated is small.

71  
72Statistical pattern recognition was used in (Mery and Filbert, 2002) to classify castings defects. In  
73statistical pattern recognition the classification is performed using the concept of similarity: similar

74 patterns are assigned to the same class (Newman and Jain, 1994). Five classifiers were used by  
75 Mery: Linear classifier, Threshold classifier, Nearest neighbour classifier, Mahalanobis classifier  
76 and Bayes classifier. For each detected region, 71 features were extracted to characterise the poten-  
77 tial flaws. The selection of the relevant features is made using sequential forward selection. Accord-  
78 ing to the authors, best performances were achieved with 100% of true defect classification and  
79 10% of false alarms rate for the Bayes classifier, and 94.7% of true defects classification and 4.1%  
80 of false alarms respectively for the threshold classifier. However, the method is validated on only  
81 76 defects.

82  
83 A flexible inspection system called ISAR (Intelligent System for Automated Radioscopy) is presen-  
84 ted in (Fuchs, 2006). Defect classification is here performed by individual evaluation of character-  
85 istic features of the segmented defect area, such as size, contrast, elongation. A special type of de-  
86 fect classification was implemented by (Wenzel, 2003). Fourier coefficients are computed in order  
87 to get translation- and rotation-invariant defect features. Those features can be assigned to a certain  
88 group, each representing a certain proper structure of the image. By computing the minimum dis-  
89 tance of an actual defect region to each of the groups, artefacts generated at regular object structures  
90 may be recognized. Our study aims to improve the ISAR system performance, especially concern-  
91 ing the false alarm rate, by introducing data fusion of objects features.

92  
93 Data fusion theory was already applied by the authors in building a classification system, where the  
94 so-called confidence levels are assigned by a source to each detected object. The fusion between  
95 confidence levels (or mass functions) is based on Dempster-Shafer theory (DS) and the final de-  
96 cision on the sample acceptability was done by a threshold in the confidence level. This approach  
97 was used in improving the detection of weld defects in (Kaftandjian et al, 2003), and in castings in-  
98 spection (Lecomte et al., 2006). The obtained results showed a more precise and reliable decision  
99 by using the data fusion approach. However, the supervision of the expert was necessary to assign  
100 the confidence levels.

101  
102 In the present work, a new mass value attribution procedure is presented, which is completely auto-  
103 matic, so that the expert supervision is no more necessary. Moreover, eleven features were com-  
104 puted for each segmented structure in the casting (instead of four features in the previous study).  
105 Each feature is considered as a source of information. By using data fusion, the aim is to build an  
106 automatic classification system. To be able to fuse different sources, a common space where values  
107 are comparable is necessary. An automatic method is introduced to convert from the space of fea-  
108 ture values into the mass value space.

109

110The rest of this paper is organized as follows. Section 2 introduces Dempster-Shafer data fusion  
111theory. Sections 3 and 4 present the proposed classification method and the performance measures.  
112Afterwards experimental results are presented in section 5, discussed in section 6 before  
113concluding.

114

## 1152. Dempster-Shafer data fusion theory

116

117Dempster-Shafer (DS) evidence theory was developed as an attempt to overcome the limitation of  
118conventional probability theory by handling uncertain, imprecise and incomplete information  
119(Dempster, 1967; Shafer, 1976). It is also suited for combining information from different sources.  
120In DS theory, a fixed set of  $N$  mutually exclusive and exhaustive elements, called the frame of  
121discernment, is defined that is symbolized by  $\Theta = \{H_1, H_2, \dots, H_N\}$ . The frame of discernment  $\Theta$   
122defines the working space for the application being considered since it consists of all propositions  
123for which the information sources can provide evidence through using the so-called mass values.  
124Information sources can also distribute mass values on subsets of the frame of discernment,  $A_i \in 2^\Theta$ .  
125Here,  $A_i$  designates a single hypothesis  $H_i$  or union of simple hypotheses (composite hypotheses).  
126An information source assigns mass values only to those hypotheses, for which it has direct  
127evidence. That is, if an information source can not distinguish between two propositions  $A_i$  and  $A_j$ , it  
128assigns a mass value to the set including both propositions ( $A_i \cup A_j$ ). This point is precisely the  
129reason for us to choose the DS theory because it reflects the hesitation between two hypotheses.

130Obtaining the mass distribution or function  $m(A_i)$  ( $0 \leq m(A_i) \leq 1$ ) is the most important step since it  
131represents the knowledge about the current application as well as the uncertainty and imprecision  
132incorporated in the selected information source. The mass distribution for all the hypotheses has to  
133fulfil the following conditions:

$$134 \quad m(\emptyset) = 0$$

$$135 \quad \sum_{A_i \in 2^\Theta} m(A_i) = 1 \quad (1)$$

136Mass distributions  $m_1, m_2$  from two different information sources are combined with Dempster's  
137orthogonal rule. The result is a new distribution,  $m = m_1 \oplus m_2$ , which carries the joint information  
138provided by the two sources:

$$139 \quad m(A_i) = (1 - K)^{-1} \times \sum_{A_p \cap A_q = A_i} m_1(A_p) m_2(A_q) \quad (2)$$

140where

$$141 \quad K = \sum_{A_p \cap A_q = \emptyset} m_1(A_p) m_2(A_q)$$

142  $K$  is often interpreted as a measure of conflict between the two sources and introduced in Equation  
143(2) as a normalization factor. The larger the  $K$ , the more conflicting are the sources and the less  
144sense makes their combination. As a consequence some authors, Smets in particular (Smets, 1990),  
145require the use of the Dempster combination rule without normalisation. However, in this case, the  
146relation (1) is no more verified. In our case, the normalised rule is used, and the  $K$  value is taken  
147into account in the decision process.

148In case of  $N$  different information sources  $B_1, B_2 \dots B_N$ , the DS rule is:

$$149 \quad m(A_i) = \frac{\sum_{B_1 \cap B_2 \cap \dots \cap B_N = A_i} m_1(B_1) m_2(B_2) \dots m_N(B_N)}{1 - K} \quad (3)$$

150where

$$151 \quad K = \sum_{B_1 \cap B_2 \cap \dots \cap B_N = \emptyset} m_1(B_1) m_2(B_2) \dots m_N(B_N) < 1$$

152From a mass distribution, two functions can be evaluated that characterize the uncertainty about the  
153hypothesis  $A_i$ . The belief function  $Bel$  measures the minimum uncertainty value about  $A_i$  whereas  
154plausibility  $Pls$  reflects the maximum uncertainty value about this hypothesis. These two measures  
155span an uncertainty interval  $[Bel(A_i), Pls(A_i)]$ , which is called “belief interval”. The length of this  
156interval gives a measurement of imprecision about the uncertainty value. Belief and plausibility  
157functions are defined from the set of symbolic set of classes  $2^\Theta$  to the unit interval  $[0, 1]$ :

$$158 \quad Bel(A_i) = \sum_{A_j \subseteq A_i} m(A_j)$$

$$159 \quad (4)$$

$$160 \quad Pls(A_i) = \sum_{A_j \cap A_i \neq \emptyset} m(A_j) \quad (5)$$

161These measures have been sometimes referred to as lower and upper probability functions.

162If only two hypotheses are used, the credibility of each single hypothesis equals its mass value, and  
163the credibility of the combined hypotheses (which represents the ignorance value) equals 1. This is  
164the case for us when considering our frame of discernment: hypothesis  $H_1$  (“this object is a true  
165defect TD”), hypothesis  $H_2$  (“this object is not a true defect”, i.e. it is a false defect FD), and the  
166ignorance is represented by the combined hypothesis ( $H_1 \cup H_2 = H_3$ ).

167Once data are combined, a decision rule must be applied. In our case, a threshold is applied on the  
168mass value corresponding to the  $H_1$  hypothesis. In the modelling step, specific care is taken in order  
169to avoid conflicts between sources.

### 1703. Proposed approach

171  
172The proposed new approach has the goal to automatically classify a detected object into either a TD  
173or a FD, without expert supervision. This is done by analysing the features values extracted from  
174each detected object, assigning to it a confidence level (or mass) and then combine mass values ob-  
175tained from different features.

176  
177In the following, the term “object” refers to the result of the segmentation process, each object be-  
178ing potentially either a true or a false defect. Both the ISAR system and our new classifier are based  
179on the same segmentation stage, so that the pre-processing does not influence the results (pre pro-  
180cessing is explained in section 5.1).

181  
182The proposed approach consists of a learning stage, and a classification stage, which are detailed  
183hereafter.

184  
185The learning is done on a population of known and manually classified objects. For those objects, a  
186set of features is computed, and the histogram of each feature is built. From these histograms, re-  
187gions of confidence are automatically defined with fuzzy transitions and masses are attributed, de-  
188pending on the true defects proportion in the region. The complete procedure is detailed below.

189  
190After this learning stage, for each object, a set of mass values is computed for each of its features. It  
191is then possible to consider each feature as a source of information about the set of hypotheses, and  
192combine them. Final decision is obtained by a threshold applied on the combined mass associated to  
193the hypothesis “defect” ( $H_1$ ).

### 194 1953.1 Learning

#### 196 1973.1.1 Estimation of true defect proportion

198  
199For a population of known detected objects (TD and FD), the set of values taken by a feature can be  
200represented in the form of a spatial repartition. The histogram is used to divide this spatial reparti-  
201tion into several regions of confidence. To illustrate the method, we use an individual normal dis-

202tributed random variable  $x$  and two sets: set A representing the class of true defects TD, and set B  
 203representing the class of false defects FD (see figure 1). An instance of  $x$  can be classified into A or  
 204B.  $x$  can take any value of the codomain of a defect feature. This example will be used for illustra-  
 205tion in each step.

206  
 207The histogram of feature distribution is plotted in figure 2 where a number of 100 intervals was  
 208chosen in the feature values range in order to have a precise estimation.

209  
 210Let  $I$  be the set of intervals forming the feature's histogram. For each interval  $i \in I$ , the percentage  
 211of instances of the class A present in this region is calculated using the following function:

$$212 \quad P_{A,B}(i) = \frac{h_A(i)}{h_A(i) + h_B(i)}$$

213  
 214  $h_A(i)$  represents the number of instances of A inside  $i$ .  
 215  $h_B(i)$  represents the number of instances of B inside  $i$ .

216  
 217**3.1.2 Elaboration of regions of confidence and associated mass functions**

218  
 219At first, each interval  $i$  is considered as a region, and a set of mass functions is attributed to this in-  
 220terval:

$$221 \quad m(H_1) = P_{A,B}(i)$$

$$222 \quad m(H_2) = 0$$

$$223 \quad m(H_3) = 1 - P_{A,B}(i)$$

224  
 225  
 226  $m(H_1)$  is the mass assigned to the hypothesis: instance of A or TD.

227  
 228 No mass is affected to the hypothesis  $H_2$  (instance of B) to ensure that no conflict occurs between  
 229 two sources of information in the fusion process.

230  $m(H_3)$  is the mass assigned to the hypothesis ignorance (instance of A or B).

231  
 232 Secondly, subsequent intervals are congregated to form a region of confidence. To be able to merge  
 233 two intervals, the variation of the function  $P_{A,B}$  is computed. Let  $\mathcal{P}_{A,B}(i)$  be the obtained function  
 234 (see figure 3).

235  
 236 If the difference between the  $P_{A,B}$  values  $|P_{A,B} - P_{A,B}|$  for two adjacent regions is less than a fixed threshold  
 237 T [first constraint], these two regions are merged and they will have in this case the same mass val-  
 238 ues:

239  
 240 If  $|P_{A,B} - P_{A,B}| < T$  : Region  $i$  is merged with region  $i+1$ .

241  
 242 Applying this method shows that the first constraint is not completely effective alone because some  
 243 regions do not contain enough points to be considered as significant. For this reason a second con-  
 244 straint on the number of points existing in each region is imposed: a region should contain at least a  
 245 certain percentage of points [second constraint] to be considered as having enough significance. Ap-  
 246 plying the merging procedure and ensuring enough points in each region, the obtained result is  
 247 shown in Figure 4, where three regions of confidence are found. The  $m(H_i)$  is specified for each re-  
 248 gion.

249  
 250 **3.1.3 Definition of fuzzy transitions between regions**

251  
 252 It can be observed in figure 4 that the mass value from one region to another may change greatly.  
 253 Therefore it is important to be able to represent the transition between two regions of the graph in a  
 254 continuous manner. The theory of fuzzy sets is used to solve this issue, by building membership  
 255 functions which allow the continuous passage from one region to another (Kaftandjian et al., 2003,  
 256 Kaftandjian et al., 2005).

257  
 258 Let  $f_k$  be a source of information used to classify an object in class A or B according to its feature  $x$ ,  
 259 and  $R_1, R_2, \dots, R_s$  be the set of regions of confidence.  $k=1\dots N$  is the number of features considered  
 260 as information sources.

261 Each object has a set of degrees of membership  $\mu_i$  ( $i=1\dots s$ ) to each region  $i$  with:

262  
 263 
$$\sum_{i=1}^s \mu_i(x) = 1$$

264 Figure 5 illustrates the fuzzy transitions obtained for the regions of figure 4. A fuzzy set is defined  
 265 and a membership function (classical trapezoidal shape) is built for each region. The slope of the  
 266 membership function is chosen to be proportional to the difference between the mass values of the  
 267 adjacent regions.

## 2693.2 Classification

270

### 2713.2.1 Mass values attribution

272

273The final mass attributed by a source  $f_k$  to the object from its feature  $x$  is calculated by weighting  
274the mass values of each region  $i$  by its degree of membership  $\mu_i$ :

$$275 \quad m(\text{object} \in H_1) = m(x / f_k) = \sum_{i=1}^s \mu_i(x) \times m(R_i) \quad (6)$$

276The final mass function obtained for the regions of figure 4 is represented in figure 5. The mass as-  
277signed to hypothesis  $H_3$  is the complement to 1 of the mass assigned to  $H_1$ .

278

$$279 \quad m(\text{object} \in H_3) = 1 - m(\text{object} \in H_1)$$

280

281Afterwards the fusion process of different sources takes place.

282

### 2833.2.2 Data fusion

284

285The Dempster-Shafer rule is used to combine each two sources among  $N$  (see eq. 2), three most op-  
286timistic sources for each object and the  $N$  sources (see eq. 3).

287

288Statistical fusion methods are also used by introducing: the mean mass value and the median mass  
289value.

290

291For an object, the mean mass is the mean value of all the mass values corresponding to this object.

292It gives a sort of equal influence to the different opinions.

293

294The median mass is the median value of all the mass values corresponding to an object. It allows to  
295introduce a majority decision for the cases where one or few sources are not in accordance with the  
296others.

297

### 2983.2.3 Decision

299

300To classify an object using the information source  $f_k$ , a threshold  $S$  is applied on its mass value  
301 $m(H_i)$  (which represents here the credibility of the hypothesis as said in section 2).

302The object is classified as:

303- a defect if  $m(H_I) \geq S$ .

304- unknown (defect or not) if  $m(H_I) < S$ .

305

#### 3064. Performance measurements and choice of optimal sources

307

308 Classification rates are introduced to measure the performance of the different combinations of  
309 information sources. These measures are calculated for different thresholds  $S$  applied on the mass  
310 values. For a certain threshold  $S$ , an object with mass  $m(H_I)$  is correctly classified if the source  
311 decision is in accordance with the human decision, this means:

312- the object is a true defect (human decision) and its corresponding mass value is higher or equal to  
313 the threshold  $m(H_I) \geq S$ .

314- or it is a false defect (human decision) and its mass is lower than the threshold  $m(H_I) < S$ .

315

316 We introduce four measures:

317

318- Correct decisions rate (PCD):

319 
$$\text{PCD} = \frac{\text{number of true defects correctly classified} + \text{false defects correctly classified}}{\text{total number of true defects and false defects}}$$

320

321- True Defects detection rate (PTD):

322 
$$\text{PTD} = \frac{\text{number of true defects correctly classified}}{\text{total number of true defects}}$$

323

324- False Defects detection rate (PFD):

325

326 
$$\text{PFD} = \frac{\text{number of false defects correctly classified}}{\text{total number of false defects}}$$

327

328

329

330- Overall detection rate R:

331 
$$\text{R} = \frac{a \cdot \text{PCD} + b \cdot \text{PTD} + c \cdot \text{PFD}}{a + b + c}$$

332

333The rate R was introduced in such a way to give the user the possibility to attribute more import-  
334ance to the correct classification of either TD or FD. Usually more importance is given to TD detec-  
335tion.

336  
337Using these measures allows comparing the final performance of the method with the existing in-  
338spection system ISAR introduced in the next section.

339  
**3405. Experimental results**

341  
342This section is devoted to present the practical application of our new method. The pre-processing  
343stage and features are described first, then the data base is presented and analysis of the results after  
344data fusion is finally detailed.

345  
**346 5.1 Pre-processing stage and selected features**

347  
348ISAR is a fully automatic system for radioscopic quality control currently in use in the production  
349of castings. Within the database used in the present paper, evaluation is done by using a reference  
350image, because thus a higher sensitivity to small defects is possible (Fuchs et al, 2006). The data  
351processing can be separated into five stages: registration, calibration, image processing, fault  
352segmentation, and quality assessment.

353 • Registration  
354Each inspected part that is processed during inline testing is not necessarily measured at exactly the  
355same position. Due to mechanical instabilities, long-term changes of the moulds and slight  
356differences between the moulds in use, an automatic registration step is needed, where the current  
357test image is registrated to the appropriate reference image using an affine transformation.

358 • Calibration  
359In order to allow for quantitative measurement of the depth of faults, i.e. to extract their length in  
360direction of the penetrating radiation, a physical calibration object is evaluated. The necessary  
361characteristic line is created using a step-wedge made out of the same material as the part and with  
362typically 10 to 15 steps covering the maximum thickness of the objects to be tested.

363 • Image processing  
364The reference image is subtracted from the current image. Thereby, all structures that do not appear  
365within the fault-free reference part will become clearly visible, that is to say potential defects. In  
366case of improper registration results however, edges due to the object structure may remain in the  
367subtraction image, which is one cause of false defects.

368 • Segmentation  
369The segmentation step is to assemble the single suspicious pixels into well defined fault regions  
370appropriately. The analysis is based on the subtraction image. As result of this operation, we will  
371have for each detected region several features that describe this object and which will be used for  
372defect classification. Each detected object in this stage is classified by ISAR as true defect or false  
373defect, and our new classification system starts with the same features.

374 • Quality assessment  
375Quality criteria are applied by ISAR to the identified defects, which are defined by the end user for  
376a final decision quality of the part (e.g. acceptance based on size or shape of defects). This post-  
377processing was not included in our system.

378  
 379The features which are selected are described as following:  
 380 • Area [mm<sup>2</sup>]: is defined as the number of pixels of the detected object multiplied by the size of  
 381 the pixel defined by the acquisition geometry.  
 382 • Depth [mm]: the main objective of using the length image is to precise the size of the defect in  
 383 X-ray penetration direction. Depth is defined as the mean value of the pixels of the detected  
 384 object.  
 385 • Volume [mm<sup>3</sup>]: is calculated using the area occupied by this object multiplied by its depth.  
 386 • MaxDepth [mm]: is the highest value inside the defect region.  
 387 • DepthDev: is equal to standard deviation of the defect's depth in the defect region.  
 388 • InOutContrast [mm]: is also measured on the length image and is equal to  
 389 
$$\text{InOutContrast} = |\text{mean}(\text{defect region}) - \text{mean}(\text{surrounding region})|$$
  
 390 • MaxElongation [mm]: is defined as the diameter of the minimum covering circle of the  
 391 detected object.  
 392 • Diff1stDev: This feature represents the difference, inside the defect region, of the mean  
 393 gradient value within the reference image to the mean gradient value within the current  
 394 image.  
 395 • InOutContrastGV: is the same as InOutContrast but in grey values instead of mm.  
 396 • Thickness [mm]: is the thickness of the part inside the defect region.  
 397 • Depth2Thickness: is the ratio between the depth of the object and the thickness of the part in  
 398 the corresponding region.  
 399

#### 4005.2 Data base description

401  
 402In this work a database is extracted from industrial images. The detected objects, also called poten-  
 403tial defects, are classified manually into TD and FD.

404  
 405The database is formed of 597 objects. It contains 382 true defects including oxides, gas voids and  
 406porosities and 215 false defects.

407  
 408The database was divided into two parts:  
 409-Learning database: formed of 115 false defects and of 243 true defects.  
 410-Testing database: formed of 100 false defects and 139 true defects.

411  
 412  
 413

#### 4145.3 Application: Learning and classification

415  
 416Spatial repartitions of four features (Area, Depth, InOutcontrast and Thickness) corresponding to  
 417the objects of the learning database are presented in figure 6. All the other features present their own  
 418spatial repartition with some overlapping between TD and FD classes. The Area graph exhibits the

419 better discrimination between classes with true defects mainly having small areas, although some  
420 false defects also.

421

422

423 First the estimations of the features histograms take place. Figure 7 illustrates the histograms of the  
424 four features represented in figure 6. As already mentioned, overlapping between TD and FD  
425 classes is important.

426

427 Using the feature's histogram the regions of confidence are built. The results for the four previous  
428 features are shown in figure 8 where different number of confidence regions are identified. The val-  
429 ues on the right correspond to the mass values  $m(H_i)$  attributed for each point falling into the corres-  
430 ponding region. In addition, membership functions were built for each feature (see figure 9).

431

432 Using the estimated regions of confidence, their corresponding mass and membership functions, the  
433 single mass values of the learning database objects are computed (eq. 6). Mass functions obtained  
434 for the four previous features are also shown in figure 9.

435

436 Afterwards the fusion process takes place. As an example, the histogram of mass values obtained  
437 from the combination of Area and InOutContrast is presented in figure 10. It can be observed that  
438 the translation from feature values to mass values results into an easier separation between TD and  
439 FD. Moreover, the fact that the mass values are included between 0 and 1 facilitates the interpreta-  
440 tion.

441 Subsequently, the classification of the learning database objects using different thresholds on the  
442 mass values given by each source (single and combined sources) is performed.

443

#### 444 **5.4 Performance results: Learning database**

445

446 The performance measures for ISAR on the learning database are: PTD=0.978, PFD=0.723 and  
447 R=0.928 (with  $a=c=1$ ,  $b=5$ ). This choice of parameters  $a$ ,  $b$  and  $c$  allows to give more importance to  
448 the classification of true defects, which is an industrial requirement.

449 Individual features performance after masses attribution is shown in Table 1. As it can be noticed,  
450 some of the features are better than the others (Area in particular, as was visible from its histogram),  
451 nevertheless, all of them will be kept for fusion due to the fact that they could bring some informa-  
452 tion.

453

454 ISAR has, in the tested configuration an excellent performance in the detection of true defects but it  
455 has a false detection rate of about 28 %. After the data fusion process, PTD, PFD and R are meas-  
456 ured for all the single and combined sources (total of 70 sources). Several combined sources give  
457 better classification rates compared to the original ISAR decision (see figure 11). All these combin-  
458 ations present better PTD, PFD and R than ISAR. Grouped into a set  $\Sigma_L$ , they are then used as  
459 sources of information to classify the objects of the testing database.

460  
461 The best classification performance as measured by the overall rate R is obtained by the DS com-  
462 bination of the features Area and Depth with 96%, for a threshold on the mass values  $S=0.8$ , instead  
463 of 92% for ISAR. Table 2 gives the results obtained for PTD and PFD of this combination which  
464 are all higher than ISAR's measures.

465  
466 When considering the performance of true defects classification only, the best source is the DS  
467 combination of Area and Thickness with PTD of 99%, for the threshold  $S=0.9$ , instead of 97% for  
468 ISAR. Here also PFD and R are both higher than ISAR's measures (see table 3).

469  
470 The best performance for false defects classification only is obtained by the mean mass with a PFD  
471 of 98 %, for the threshold  $S=0.7$ , which is 28% better than ISAR's PFD. However, PTD is in this  
472 case lower than ISAR's PTD which is not acceptable for this application (see table 4), therefore the  
473 mean mass is not presented in figure 11. Nevertheless, this source will be considered for testing.

#### 474 475 **5.5 Validation: Testing database**

476  
477 The testing database is formed of 139 true defects and 100 false defects.

478  
479 ISAR shows similar performance on the testing database with: PTD=0.980, PFD=0.725 and  
480 R=0.930.

481  
482 After the learning step, the set of sources  $\Sigma_L$  are used to automatically attribute mass values for  
483 the new objects of the testing database. Afterwards data fusion and classification take place.

484  
485 All combined sources improve with an important percentage the classification of false defects and  
486 also are equal or even exceed ISAR performance on true defects detection (see figure 11 (b)). In the  
487 learning database, the optimal overall source was the fusion between Area and Depth. On the testing  
488 database, this combination presents a rate R of 96.37 % (for the threshold  $S=0.8$ ), instead of 93.07%

489for ISAR, with a very small difference from the optimal overall source on the testing database  
490which is the fusion between Area and MaxElongation with  $R=96.86\%$  (see table 5).

491  
492 The optimal combination for the detection of true defects in the learning database (DS combination  
493of Area and Thickness) is still the best combination for the detection of true defects in the testing  
494database with a PTD of 0.99% for the threshold  $S=0.9$  (see table 6). The improvement in the detec-  
495tion of true defects is 1 % relatively to ISAR, and 12 % in the detection of false defects. The optim-  
496al combination for the detection of false defects in case of testing database is still the mean mass  
497value for the threshold  $S=0.7$  (see table 7). The improvement in the detection of false defects is 25.8  
498% relatively to ISAR, but a loss of almost 4.5 % in the percentage of true defects detection occurs.  
499To illustrate two cases where ISAR fails, while data fusion gives the good result, two objects are  
500chosen. Figure 12a shows a false defect classified as a true one, and figure 12b shows a true defect  
501classified as a false (the object appears brighter than the background inside the rectangular zone).

502

503

504

## 5056. Discussion

506

507The first point to emphasize is that the results obtained during the testing phase confirm the  
508robustness of the method. All sources selected during learning keep their high performance during  
509testing.

510

511In our previous work (Kaftandjian et al., 2003) only 4 features were selected as best candidates to  
512discriminate between TD and FD. This selection of features (contrast to noise ratio, area, elongation  
513and position) was done by the expert based on his knowledge of X-ray inspection. The drawback  
514was to require the expert's supervision during learning. Here the approach is completely automated  
515and all the features are considered without a priori information. The obtained performance cannot  
516be compared with (Kaftandjian et al., 2003) due to the fact that the application was not the same.

517When considering which features give the best classification rates, one can be surprised that the  
518contrast feature is not among the best ones, although contrast is usually considered as having the  
519main physical meaning. Indeed it is impossible to justify a posteriori why some features were better  
520than the others when combined. This is due to the fact that our knowledge is only valid for TD, and  
521the classification of FD cannot be based on a physical basis. This is precisely the difficulty of FD  
522detection and the reason why even the expert find them difficult to classify. Therefore an important  
523advantage of our method is that it adapts itself to all the objective data that are measured on the true

524and false defects, without any a priori knowledge.

525

526Another fact to mention is that the developed method is general in the sense that it can work for any  
527inspection method. The translation from the feature space to the mass values space allows to  
528combine any information sources independently of their origin (ultrasonic data, radiographic  
529data...).

530

## 531 **7. Conclusion**

532

533A new implemented method able to automatically detect regions of confidence from a feature's  
534histogram and to affect degrees of confidence (mass values) to these regions was presented. The  
535automatic learning process is based on a classified database. The method consists in translating a  
536feature value into a confidence or mass value. These mass values are used within the fusion process.  
537The method was applied on a learning database and validated on a testing database for a casting  
538application. As could be shown, several feature combinations are exceeding the current ISAR  
539decision. Using the data fusion theory, while conserving at least the 98 % percentage of  
540classification of TD given by ISAR, an improvement from 72 % to 92 % (see table 4) in the  
541percentage of FD classification was achieved. An improvement of 3.8% of the overall rate R is  
542obtained.

543

544Some combinations can optimize even better either the TD or the FD classification. A PTD rate of  
54599% and PFD of 98.3% was obtained separately. Further work is needed to find an optimal decision  
546rule which would associate the best TD and FD classification rates.

547

548In a future work, a study of the robustness of this method for another application will be done.  
549Further investigation will also be conducted to measure the influence of the constraints on the  
550derivative variation and on the percentage of points inside the regions of confidence on the results.

551

## 552 **Acknowledgements**

553

554This work was co-financed by the European Union and the Free State of Bavaria within the project  
555ZeLuR.

556

557

## 558 **References**

559  
560Bonser G., Lawson S.W., "Defect detection in partially complete SAW and TIG welds using on-  
561line radioscopy and image processing". Proceedings of SPIE International Symposium on Non  
562Destructive Evaluation Techniques for Aging Infrastructure and Manufacturing, pp. 231-239, 1998,  
563San Antonio Texas USA.

564  
565Dempster A., "Upper and lower probabilities induced by multivalued mapping", Annals of  
566Mathematical Statistics, 38, pp. 325-339, 1967.

567  
568Dobrazanski L.A., Krupinski M., and Sokolowski J.H., "Computer aided classification of flaws  
569occurred during casting of aluminum". Journal of Materials Processing Technology, Vol. 167,  
570Issues 2-3, pp. 456-462, August 2005.

571  
572Fuchs T., Hassler U., Huetten U., and Wenzel T., "A new system for fully automatic inspection of  
573digital flat-panel detector radiographs of aluminium castings". Proceedings of 9th European  
574Conference on Non-Destructive Testing (ECNDT), Sept. 25 – 29, 2006, Berlin Germany.

575  
576Hernandez S., Saez D., Mery D., Silva R., and Sequeira M., "Automated defect detection in  
577aluminum castings and welds using neuro-fuzzy classifiers". 16th World Conference on NDT, Aug  
57830 - Sep 3, 2004, Montreal Canada.

579  
580Jang J., "Anfis: Adaptive-network-based fuzzy inference system" IEEE Transactions on Systems,  
581Man, and Cybernetics, Vol. 23, No. 3, pp. 665-684, 1993.

582  
583Kaftandjian V., Dupuis O., Babot, D. and Zhu Y., "Uncertainty modeling using Dempster-Shafer  
584theory for improving detection of weld defects". Pattern Recognition Letters, Volume 24, pp. 547-  
585564, 2003.

586**Kaftandjian V.**, Zhu Y.M., Dupuis O., Babot D., The combined use of the Evidence Theory and Fuzzy  
587Logic for improving multimodal Non Destructive Testing systems, *IEEE Trans. On Instrumentation &*  
588*Measurement*, 2005, vol.54, n°5, p.1968-1977.

589  
590Lawson S.W. and Parker G.A., "Intelligent segmentation of industrial radiographic images using  
591neural networks". Machine Vision Applications and Systems Integration III, Proceedings of SPIE,  
592pp. 245-255, November 1994.

593

594Lecomte G., Kaftandjian V., and Cendre E., "Combination of information from several x-ray  
595images for improving defect detection performances-application to castings inspection".  
596Proceedings, 9th European Conference on Non-Destructive Testing (ECNDT), Sept. 25 – 29, 2006,  
597Berlin Germany.

598  
599Mery D. and Filbert D., "Classification of potential defects in the automatic inspection of aluminum  
600castings using statistical pattern recognition". In Proceedings of the 8th European Conference on  
601Non-Destructive Testing (ECNDT 2002), June 2002, Barcelona Spain.

602  
603Newman T. and Jain A., "Bidirectional template-matching for 3d cad-based inspection".  
604Proceedings of SPIE Machine Vision Applications in Industrial Inspection II, San Jose, pp. 257-  
605265, February 1994.

606  
607Shafer G., "A mathematical theory of evidence", Princeton University Press, Princeton, pp. 297,  
6081976.

609Smets P., "The combination of evidence in the Transferable Belief Model", IEEE trans. Pattern  
610Analysis & Machine Intelligence, vol 12, n°5, pp. 447-458, 1990.

611  
612  
613Theis G. and Kahrs T., "Fully automatic X-ray inspection of aluminum wheels". Proceedings of the  
6148th European Conference on Non-Destructive Testing (ECNDT 2002), June 2002, Barcelona Spain.

615  
616Wenzel T. and Hanke R., "Texture analysis for classification of cast defects in x-ray images". 2<sup>nd</sup>  
617workshop "NDT in progress" International Meeting of NDT Experts, October 6-8, 2003, Prague,  
618Czech Republic.

619  
620  
621  
622  
623  
624  
625  
626  
627

628**Captions**

629

630**Fig. 1:**

631Two normal distributions where a star point represents a member of the class A and a cross point  
632represents a member of class B.

633**Fig. 2:**

634Feature's histogram of class A (full line) and class B (dotted line).

635**Fig. 3:**

636Illustration of  $P_{A,B}(i)$  (full line) and  $\approx P_{A,B}(i)$  (dotted line) corresponding to the histogram of fig-  
637ure 2.

638**Fig. 4:**

639Regions of confidence after merging and their corresponding mass values  $m(H_i)$ .

640**Fig. 5:**

641Membership functions corresponding to the regions of figure 4 (full line) and mass functions (dot-  
642ted line).

643**Fig. 6:**

644Spatial repartition of True and False defects for Area (high left), Depth (high right), InOutContrast  
645(low left) and Thickness (low right)

646**Fig. 7:**

647Histograms for Area (high left), Depth (high right), InOutContrast (low left) and Thickness (low  
648right).

649**Fig. 8:**

650Regions of confidence and corresponding mass values  $m(H_i)$  for Area (high left), Depth (high  
651right), InOutContrast (low left) and Thickness (low right).

652**Fig. 9:**

653Membership functions (solid line) and mass functions (dotted line) corresponding to the confidence  
654regions of Area (high left), Depth (high right), InOutContrast (low left) and Thickness (low right)

655**Fig. 10:**

656Histogram of mass values obtained from the combination of Area and InOutContrast. The majority  
657of true defects have mass values close to one.

658**Fig. 11:**

659Classification rates obtained on the learning database (a) and on the testing database (b). The selec-  
660ted sources are those giving a better overall classification rate than ISAR, and the same set of  
661sources are still better on the testing database

662

663**Fig. 12:**

664 False defect classified as a true one by ISAR (a) and true defect classified as false (b). In those two  
665 cases, data fusion give the good result. The true or false defects appear as a brighter zone in the  
666 rectangle

667

668 **Table 1:**

669 Performances of single features after mass attribution, in the learning database. Features are classi-  
670 fied with respect to the overall rate R

671

672 **Table 2:**

673 Combination giving the highest overall rate R in the learning database.

674

675 **Table 3:**

676 Combination giving the highest true defects detection rate PTD in the learning database.

677

678 **Table 4:**

679 Combination giving the highest false defects detection rate PFD in the learning database.

680

681 **Table 5:**

682 Combination giving the highest overall rate R in the testing database.

683

684

685 **Table 6:**

686 Combination giving the highest true defects detection rate PTD in the testing database.

687

688 **Table 7:**

689 Combination giving the highest false defects detection rate PFD in the testing database.

690

691