Weld defect classification using EM algorithm for Gaussian mixture model

M.Tredi, S.Belaifa et N.Nacereddine

Laboratoire de Traitement du signal et de l'image
Centre de recherche scientifique et technique en soudage et contrôle (CSC)
Route de dély I brahim, B.P.64 Chéraga, Alger, Algerie
m_tridi@hotmail.com
s_belaifa@hotmail.com
necereddine_naf@hotmail.com

Abstract: In this paper we present a new approach to classify weld defects from radiogram images based on Expectation Maximization (EM) algorithm. The EM is one most algorithm used in statistical pattern recognition. However, it is not yet commonly exploited in the classification of radiogram images in NDT. The approach to detecting weld defects follows a general pattern recognition scheme based on three steps: segmentation, feature extraction and classification. That is in our case, 1) the defects are segmented by using an histogram thresholding; 2) the feature vector of the defect are extracted by using geometrical parameters; and 3) the EM and FCMI algorithms are used for classification. The results show that the proposed algorithm performs much better than FCMI algorithm, and demonstrate that it is efficient for weld defects classification of radiogram images.

Keywords: EM algorithm, FCMI algorithm, Gaussian mixture, Radiogram images, Weld defect.

1 Introduction

To evaluate the quality of a weld joint, many methods of non-destructive testing (NDT) are used. The external defects are generally detected by a visual examination with the use of the measuring instruments while the interior defects are detected by physical methods: Magnetic, X-ray, Gamma-ray, ultrasounds, etc.

Radiography testing is one of major NDT methods to examine weld defects. In most applications, a radiographic weld image is produced by permitting an X-ray or gamma-ray source to penetrate the weld component and expose a photographic film. This method is used for inspecting several types of welds assemblies such as pipe-lines, boilers, pressure vessels etc. Inspected zones may present multifarious defects such as porosity, inclusion, crack, lack of penetration, lack of fusion etc. This radiogram is examined by radiography interpreters whose the task consist in detect, recognize and quantify eventual defects, but the radiogram quality, the welding over-thickness, the bad contrast, the noise and the weak sizes of defects make difficult their job. The defect quantification in these conditions is submitted to human judgement and subjective considerations, such as, capabilities and experience of the interpreter because, it takes times to train a film interpreter. The manual interpretation process is often subjective, inconsistent, labour intensive, and biased prone. It is thus desirable to develop some forms of computer-aided systems to assist the human interpreter in evaluating the quality of welding joints. This involves the digitalization of film radiography and the development of algorithm to extract welds and to identify flaws in them, (Nacereddine & al., 2004). The data processing sequences of radiogram images is given in the figure 1.

Many works have been proposed for the classification and pattern recognition of weld defects (Belaifa & al., 2004), (Mery & al., 2003), (Nacereddine & al., 2004) and (Warren Liao & al., 2003 ). (Nacereddine & al., 2004) use the artificial neural networks trained by the back propagation, to the primary classification of the weld defects in radiogram images (planer defect, volumetric defect), through image analysis by geometric invariant moments. (Belaifa& al., 2004) Give the components of features vector by the geometrical parameters. In
(Mery & al., 2003), the others detect the weld defects using texture features.

In this work, we propose two algorithms of classification of the weld defects (Fuzzy C-Means Iterative algorithm: FCMI and Expectation Maximisation: EM). The first algorithm is based on the concept of distance and fuzzy logical and the second based on the statistical concept.

The rest of the paper is organized as follows. In section 2, we review the basic principles of FCMI algorithm, Gaussian mixtures and the EM algorithm. In section 3, we apply the proposed algorithms to the classification of weld defects on radiogram images. Some conclusions are given in section 4.

2 Problem formulation

2.1 The FCMI algorithm

Classification or decision by the distance is one of the most simple and the most intuitive approaches of the pattern recognition. The first motivation of such approach is to consider that a sample belongs to one cluster among the other clusters. Fuzzy-C-Means-Iterative algorithm uses the concept of fuzzy logic and distance for classification. It is given by the following algorithm [3]:

a) $M$ cluster centers are initialised in the set of samples.

b) Calculation of the euclidean distance between each sample and each cluster center

$$d_{iy} = \| y_i - y_i \|$$

(1)

c) Calculation of the memberships function $\chi$ (fuzzification)

$$\lambda_{iy} = \sum_{y=1}^{M} \left( \frac{d_{iy}}{d_{iy}} \right)^{2(\beta-1)}$$

(2)

$\beta$ is a tuning parameter which controls the degree of fuzziness in the process.

If $d_{iy} = 0$ for $l = l_0$ then $\lambda_{iy,l_0} = 1$ and ($\lambda_{iy} = 0$ for $1 \leq i \leq M$ and $i \neq l_0$).

d) The cluster centers change by using the following formula:

$$y_i = \frac{\sum_{j=1}^{M} \lambda_{ij} x_j}{\sum_{j=1}^{M} \lambda_{ij}}$$

(3)

with m: the number of samples.

e) If the new cluster centres are changed, then we move to the step b) else we continue.

f) Defuzzification: if $\lambda_{iy,l_0} = \max(\lambda_{ij})$ then $\lambda_{iy,l_0} = 1$ and $\lambda_{ij} = 0$ for $1 \leq i \leq M$, $1 \leq j \leq m$ and $i \neq l_0$. The form $x_j$ belongs to the class $l_0$.

2.2 Gaussian mixture model.

Consider mixture model with $M > 1$ components (classes) in $R^n$ for $n \geq 1$.

$$p(x) = \sum_{m=1}^{M} \alpha_m p(x_{/m})$$

(4)

Where $\alpha_m \in [0,1]$ (m=1,2,...., M) are the mixing proportions subject to

$$\sum_{m=1}^{M} \alpha_m = 1$$

(5)

For Gaussian mixtures, each component density $p(x_{/m})$ is a normal probability distribution:

$$p(x_{/m}) = \frac{1}{(2\pi)^{n/2} \det(C_m)^{1/2}} \exp \left[ -\frac{1}{2} (x - \mu_m)^T C_m^{-1} (x - \mu_m) \right]$$

(6)

where T denotes the transpose operation. Here we encapsulate these parameters into parameter vector, writing the parameters of each component as $\theta_m=(\mu_m, C_m)$ to get $\Theta = (\alpha_1, \alpha_2, ..., \alpha_m, \theta_1, \theta_2, ..., \theta_m)$ then Eq(6) can be rewritten as

$$p(x_{/\Theta}) = \sum_{m=1}^{M} \alpha_m p(x_{/\theta_m})$$

(7)

If we knew the component from which x came,
then it would be simple to determine the parameters $\Theta$ similarly, if we knew the parameter $\Theta$, we could determine the component that would be most likely to have produced $x$. The difficulty is that we know neither. However, the EM algorithm could be introduced to deal with this difficulty through the concept of missing (Zhihua & al, 2003).

2.3 The EM algorithm

Expectation maximization (EM) algorithm is a widely used class of iterative algorithms for maximum likelihood (ML) or maximum posteriori (MAP) estimation in problems with missing data. Given a set of samples $X=(x_1,x_2,...,x_k)$, the complete data set $Z=(X,Y)$ consists of the sample set $X$ and a set $Y$ of variables indicating from which component of the mixtures the sample came. Now we discuss how to estimate the parameters of the gaussian mixtures with the EM algorithm (Zhihua & al, 2003).

The EM algorithm consists of an E-step and M-step. Suppose that $^\Theta$ denotes the estimation of $\Theta$ obtained after the $t$th iteration of the algorithm. Then at the $(t+1)$th iteration, the E-step computes the expected complete data log-likelihood function

$$Q(\Theta^{(t)}) = \sum_{k=1}^{K} \sum_{m=1}^{M} \left[ \log p(m/x_k; \Theta_m) \right] P(m/x_k; \Theta^{(t)})$$

where $P(m/x_k; \Theta^{(t)})$ is a posterior probability and is computed as

$$P(m/x_k; \Theta^{(t)}) = \frac{\alpha_m^{(t)} p(x_k / \Theta_m^{(t)})}{\sum_{l=1}^{M} \alpha_l^{(t)} p(x_k / \Theta_l^{(t)})}$$

And the M-step finds the $(t+1)$th estimation $\Theta^{(t+1)}$ of $\Theta$ by maximizing $Q(\Theta^{(t)})$.

$$\alpha_m^{(t+1)} = \frac{1}{K} \sum_{k=1}^{K} P(m/x_k; \Theta^{(t)})$$

$$\mu_m^{(t+1)} = \frac{\sum_{k=1}^{K} x_k P(m/x_k; \Theta^{(t)})}{\sum_{k=1}^{K} P(m/x_k; \Theta^{(t)})}$$

$$C_m^{(t+1)} = \frac{\sum_{k=1}^{K} P(m/x_k; \Theta^{(t)}) (x_k - \mu_m^{(t+1)})^T (x_k - \mu_m^{(t+1)})}{\sum_{k=1}^{K} P(m/x_k; \Theta^{(t)})}$$

3. Application in weld defect classification

In our application we have taken data base of 72 original radiogram images of weld defects. $R$ is an original image, after segmentation, we obtain the image $f$ (see figure2,3).

![Original images(R)](image1)

![Image segmented (f)](image2)

To be able to classify a pattern (image segmented), it is essential to characterize it by the features vector. The choice of this vector is based on knowledge obtained by the expert in radiography. We can find many kinds of features vector using Zernik moments, Legendre moments, Geometric moments, Fourier coefficients, etc.

The features used in our application are the geometrical parameter or geometrical feature. This type of feature consists in characterizing an object according to vector whose elements are characteristics, such as perimeter, surface, principal direction of inertia and elongation.

$$x=[S, \alpha, L, W, P, E]^T$$

$S$: Area of defect.
$L$: The length of the defect.
$W$: The width of the defect.
$\alpha$: Principal direction of inertia calculation.
$E$: Elongation.
$P$: Perimeter

we have supposed that the number of clusters $M=4$, such as: 20 cracks, 21 lack of penetration, 17 gas inclusions and 14 oxide inclusion in the data base of 72 defects. Here are some radiogram images which represent the four defects:
3.1 FCMI method

We notice that the cluster centers represent efficiently the four classes (y1 for cracks, y2 for lack of penetration, y3 for gas inclusion, y4 for oxide inclusion), and they are different from each other. The choice of $\beta$ is important to ensure a higher classification rate by FCMI method with $(\beta = 1.09)$. 

\[
\begin{align*}
\tilde{\mu}_1 &= [52.55 \ -0.02 \ 36.97 \ 4.07 \ 63.98 \ 8.53]^T \\
\tilde{\mu}_2 &= [565.57 \ 0.29 \ 222.71 \ 25.48 \ 284.27 \ 16.90]^T \\
\tilde{\mu}_3 &= [19.21 \ 19.01 \ 6.86 \ 2.43 \ 18.10 \ 2.95]^T \\
\tilde{\mu}_4 &= [15.53 \ -16.28 \ 3.94 \ 3.18 \ 14.54 \ 1.23]^T
\end{align*}
\]

Each sample is classified in one of classes by using Bayes classifier. The following table (table 1) summarises the rates of classification for each method:

<table>
<thead>
<tr>
<th></th>
<th>Nb1</th>
<th>Nb2</th>
<th>Nb3</th>
<th>Nb4</th>
<th>Rate of classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCMI</td>
<td>22</td>
<td>19</td>
<td>19</td>
<td>12</td>
<td>88.88%</td>
</tr>
<tr>
<td>EM</td>
<td>20</td>
<td>21</td>
<td>17</td>
<td>14</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 1: Rate of classification

Nb1, Nb2, Nb3 and Nb4 are the numbers of crack, lack penetration, gas inclusion and oxide inclusion.
calculated respectively

**Conclusion**

We have described a new approach to classify the weld defect for radiogram images using EM algorithm. The EM algorithm is very sensitive to the choice of the initial values of parameters. In our case, we have used the k-means algorithm for initialization.

The main contribution is a comparison between EM and FCMI algorithms. The experimental results show that our algorithm has given better results than FCMI algorithm.

Other methods in statistical pattern recognition can be used for classification, the application of them in weld defects classification is one of our future research focuses. And also, it is recommended to increase the size of the features vector and the data base in order to identify great classes of weld defects which exist in industry.

**References**


