Visual Evidence Accumulation in Radiograph Inspection

Horace H. S. Ip

Department of Computer Science
City Polytechnic of Hong Kong
Tat Chee Avenue
Kowloon
Hong Kong

Abstract

Image features pertinent to weld defect detection and identification are extracted from a digitised radiograph image of the weld. These image features form the set of visual evidence which is brought to bear upon a set of possible defect hypotheses. The Dempster-Shafer theory is applied to combine these visual evidence and obtain a belief interval for each of the defect hypotheses. The system is capable of assessing the validity of the result of the identification by considering the degree of conflict in the body of the evidence.

1. Introduction

This paper presents an approach to weld defect identification based on the accumulation and combination of visual evidence extracted from a radiograph image of the weld. Uncertainties in machine inspection of radiographs arise as a result of: (a) uncertainty in the detection of a specific visual evidence due to possible errors in the image segmentation and feature extraction processes; (b) uncertainty in the conclusions that should be drawn from the evidence. The Dempster-Shafer theory provides a mathematical basis for combining evidence which has been brought to bear upon a set of hypotheses and for reasoning under uncertainties [1,2].

1.1 Dempster-Shafer Theory

Within the framework of the theory, the set of all possible hypotheses within the problem domain is called the frame of discernment, or Θ. Hypotheses within Θ are mutually exclusive and exhaustive. A piece of evidence can be brought to bear upon one or more subsets of Θ. Furthermore, each piece of evidence x has associated with it a mass function $m_x(H)$ which expresses the degree to which the evidence supports or refute a hypothesis H. The mass function has a range of [0,1]. The total degree of belief on a hypothesis as a result of pooling several pieces of evidence can be computed using the Dempster's rule:

$$m_{ab}(H) = K \, . \sum_{P \cap Q = H} m_a(P) . m_b(Q) \tag{1}$$

$$K = 1 - \sum_{P \cap Q = \phi} m_a(P) \, . \, m_b(Q)$$

BMVC 1991 doi:10.5244/C.5.26

where $m_{ab}(H)$ denotes the amount of belief mass assigned to hypothesis H as a result of combining two pieces of evidence, a and b; and K is a normalisation term which ensures that the total mass assigned to the focal hypotheses and Θ summed to 1, and that the mass assigned to the empty set is zero. The subsets of Θ to which the mass function assigned non-zero mass are called the focal hypotheses of the evidence.

1.2 Belief Functions and Belief Interval

A belief function denoted *Bel(H)* measures the degree to which the available evidence directly supports the hypothesis H. This is expressed as the sum of the mass assigned to H and all its subsets, ie.

$$\mathrm{Bel(H)} = \sum_{h\ \ H} \mathrm{m(h)} \tag{2}$$

The plausibility of the hypothesis H, *Pl(H)*, which is the degree to which the available evidence fails to disconfirm H can be expressed as

$$\mathrm{Pl(H)} = 1 - \mathrm{Bel}(\tilde{}\mathrm{H}) \tag{3}$$

The belief interval for a hypothesis H is given by

$$[\mathrm{Bel(H)},\ \mathrm{Pl(H)}]$$

which represents explicitly the support and plausibility of a proposition H. The belief interval [1,1] and [0,1] indicates complete certainty and complete uncertainty of the hypothesis respectively.

2. Visual Evidence Elicitation in Weld Defect Identification

A set of visual cues which can be derived from the weld radiograph and are deemed pertinent to defect identification were elicitated using a combination of knowledge acquisition techniques [3], namely, document analysis, protocol analysis and goal-decomposition method. Initially, an experienced radiographer was asked to "think aloud" while carrying out the identification task on a sample of weld radiographs, watched by two knowledge engineers. The engineers asked questions designed to clarify the radiographer's actions and his working hypotheses. The entire session was taped and analysed off-line.

Next, the radiographer was replaced by a knowledge engineer who acted as the "eyes" for the radiographer. The radiographer had to decide whether the weld in fact contained a defect, and if yes, what the defect was by asking the engineer questions concerning the visual features which could seen on the radiograph. Again, the entire session was taped and analysed off-line. The process was repeated a number of times.

In document analysis, relevant documents defining the different types of defect and their causes were studied and analysed. These documents provide information relating to the deep knowledge of the problem domain, eg. the underlying physics of the occurence of a defect and its physical properties.

The knowledge elicitation process identified the visual cues which relate a defect and the degree of the defect severity to pictorial features that can be seen or extracted from the weld radiograph. Figure 1 summarises some of the results of the elicitation process. This process identified nine salient features deemed relevant to defect identification. They are:

1. skeleton length
2. intensity differece with respect to the weld
3. shape
4. width
5. size
6. orientation
7. location with respect to weld medial axis
8. noise like
9. dark feature

Furthermore, each of these visual features has associated with it a set of possible attribute values. For example, skeleton length can be long, medium or short; shape can be elongated or circular, etc.

3. Formulation of the Identification Problem Under D-S Theory

Under D-S theory, the weld defect identification problem can be posed as follows:

(a) The frame of discernment of the problem domain consists of the set of possible weld defect hypotheses which for the purpose of this study has been restricted to five defect types, ie. {gas pore (GP), crack (CK), excess penetration (EP), lack of root fusion (LRF) and root concavity (RC)}.

(b) The set of evidence which will form the input to the identification system consists of the set of nine visual features defined above.

Our approach is to design image processing techniques for extracting visual evidence from an image of the weld radiograph. These pieces of evidence are subsequently combined using the Dempster's rule to yield a set of belief intervals for the competing defect hypotheses. A successful application of the theory depends on the solutions to the following practical issues:

Given a piece of the evidence, the identification system has to decide

(a) what are the set of focal or defect hypotheses?
(b) what is the amount of belief mass to be assigned to the focal hypotheses?
(c) what is the degree of conflict in the body of evidence presented?

The theory itself gives no indication as to how these issues should be resolved for an application. Frequently, the solutions to the first two issues are "fixed" by the "domain experts" during system design. Consequently, the rationale behind why certain belief mass is assigned to a particular hypothesis are not easily traced or lost entirely. Since the total number of possible hypotheses is the power set of Θ, it is not possible to predefine the belief mass to be assigned to each of the possible hypotheses based on expert opinions. The following sections describe our solutions to these issues.

3.1 Defect Hypothesis Generation

Here we adopt the strategy of hypothesis elimination and contend that disconfirming evidence is a better source of information than confirming evidence. For example, if we detected a circular shaped suspect defect, it is highly likely that it is not a crack defect whereas the same piece of evidence only weakly suggests that the defect may either be a gas pore, excess penetration, or metal inclusion - further measurement on the intensity characteristic of this suspect defect can help to distinguish between these possibilities. The above consideration led to the following hypothesis generation strategy:

Given a detected image feature

(a) the feature is matched against a knowledge base of defect objects and the set of defect object H which does not match the detected feature are selected

(b) the set negation of H with respect to Θ is computed and is identified as the defect hypothesis inducecd by the piece of visual evidence.

The term "does not match" here includes those defects which we know of its existence but do not have any detailed information on the features being matched. This way, we accept the possibility of ignorance and take a conservative approach of using the body of evidence. Furthermore, we grant the benefits of the doubt to those defects which we do not have detailed information about it. This strategy therefore bias towards reducing the false negative rate of the identification results. The latter is particularly important for safety critical applications. For each piece of the evidence, this approach will generate at most one focal hypothesis.

3.2 Mass Assignment

From a probablistic point of view, the mass distribution over the set of focal hypothesis induced by the evidence is related but not equivalent to the *posterior probability* of the focal hypothesis given the evidence [4,5]. For this application, it can be seen from figure 1 that the existence of a particular pictorial feature, eg. line-like object, implies the existence of a number of possible defect types, e.g. crack or lack of root fusion. Consequently, the focal hypothesis H of the evidence is in general non-singleton and consists of a disjunction of singleton defect propositions, ie. $\{h_1, h_2, \ldots, h_n\}$. Assuming that the occurrence of individual defect types are independent, the *posterior probability* of the hypothesis H, given a piece of evidence E, denoted by $P(H|E)$ is:

$$P(H|E) = \sum_{h_i \in H} P(h_i|E)$$

Furthermore, Bayes Theorem gives

$$P(h_i|E) = P(E|h_i).P(h_i)/P(E)$$

where $P(E|h_i)$, $P(E)$, $P(h_i)$ are the prior probabilities of the evidence given the hypothesis, prior probability of the evidence and of the hypothesis respectively.

Assuming all defect types are equally probable, ie. $P(h_i) = 1/N$, where N is the cardinality of the frame of discernment, a semi-empirical mass function $m_E(H)$, due to evidence E whose focal hypothesis is H, can be written as:

$$m_E(H) = M.a(E)/[M.a(E) + (N\text{-}M).b(E)] \tag{4}$$

where the parameters $a(E)$ and $b(E)$ reflect the reliability of detecting the evidence (visual feature); and M denotes the cardinality of the focal hypothesis set. The remaining mass of $(1\text{-}m_E(H))$ is assigned to Θ. Detail of the derivation of equation (4) can be found in [6].

Since at most one focal hypothesis will be generated using our hypothesis generation strategy, the associated mass function is termed a *simple mass function*. For simple mass functions, the maximum number of competing hypotheses resulting from combining M pieces of evidence is at most 2^M, including Θ. This sets the upper bound on the computational load of the reasoning process. By limiting the number of visual evidence available to the reasoning process, we can estimate the complexity and the worst case response time of the system [7].

3.3 Conflicting Evidence

Contradiction in the reasoning process can arise as a result of (a) errors in the segmentation and feature detection processes; (b) error in the hypothesis or conclusion that have been drawn from the evidence. When two pieces of conflicting evidence is combined under D-S theory, certain amount of mass may accrue in the empty set. The normalisation factor K in the Demspter's rule is designed to re-distribute this "redundant belief" among all the competing hypotheses. We contend here that, for simple mass functions, the mass accrued in the empty set is related to the degree of contradiction or inconsistency in the reasoning process and should be retained throughout the reasoning process.

The space of *unnormalised belief states* introduced in [8] showed that this new space can be mapped homomorphically onto the original Dempster's rule space. This means that we can maintain the mass accrued in the empty set througout the evidence accumulation process without losing any information concerning the degrees of support and plausibility for the competing hypotheses. After the body of evidence has been pooled, the mass remained in the empty set gives us an assessment of the degree of conflict in the evidence and hence the validity of the result of the defect identification.

4. System Overview

The identification system consists of two subsystems for image processing and feature extraction and for evidence combination respectively. The former subsystem which ran on a cellular array image processor [9] delineated the weld region from the digitised radiograph image. Features which are darker or lighter than the nominal weld intensity were subsequently enhanced and extracted by means of a series morphological filtering operations [10]. For each of these suspect objects, a set of feature measurements were made. These numerical measurements were converted into a set of symbolic descriptors by means of a set of production rules (Figure 2). These descriptors form the set of visual evidence to be used in the defect identification process.

4.1 Preliminary Results

The evidence combination subsystem combined the set of visual evidence and classified each suspect objects detected within the weld according to the belief interval computed for each element (or defect type) of Θ and the mass accrued in the empty set. The classification process works as follows: first the defect type x with the most support and plausibility is identified, if the plausibility of the defect type is less than the support mass accrued in the empty set, then the suspect object is classified as being defect x, otherwise, the system concludes that the body visual evidence presented is conflicting or inconsistent. The system at this point may be programmed to either call for human intervention or declare that the suspect object is an artefact.

A prototype system has been built to investigate the feasibility of this approach. Preliminary results indicated that for a genuine defect, the body of visual evidence derived from the image was highly consistent (with a very low or zero mass for the empty set) and yielded a high degree of support and plausibility for the corresponding defect hypothesis. On the other hand, if the suspect object was in fact an artefact of the radiograph, the resulting body of evidence was highly inconsistent and gave rise to a significant amount of mass accrued in the emtpy set. Figure 3 gives an example of the system output.

5. Conclusions

This paper presents an evidential reasoning approach to weld defect identification. Preliminary results indicated that our solutions to application issues such as hypothesis generation and mass assignment strategies are highly appropriate to an application domain where the body of evidence is uncertain and tends to weakly support a disjunction of object classes.

The identification system has the following characteristics: (a) it is a strict application of D-S theory, (b) it supports mass assignment to an arbitrary subsets of object classes within the frame of discernment, (c) the mass accrued in the empty set is used to assess the degree of conflict or inconsistency within the body of evidence.

References

[1] Dempster, A. P. *A Generalisation of Bayesian Inference*. J. Roy. Statis. Soc., Vol. B30, pp.205-246, 1968.

[2] Shafer, G. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ., 1976.

[3] Kidd, A. *Knowledge Acquisition for Expert Systems: A Practical Handbook.* Plenum Press, 1987.

[4] Yen, J. *GERTIS: A Dempster-Shafer Approach to Diagnosing Hierarchical Hypotheses*. Comm. ACM., Vol. 32, No. 5, pp.573-585, 1989.

[5] Pearl, J. *Bayesian and Belief Function Formalisms for Evidential Reasoning: A Conceptual Analysis*. in Readings in Uncertain Reasoning, Ed. G. Shafer and J. Pearl, pp. 540-574, Morgan Kaufmann, USA, 1990.

[6] Ip, H. H. S. *Application of a Theory of Evidence in Knowledge-Based Pattern Recognition*. Proc. NCIT'91, pp. 98-110, Penang, Malaysia, 1991.

[7] Ip, H. H. S. and Bell, M.*An Evidence Reasoning Scheme for Intercept Recognition*. Technical Memorandum, C3119-TM-011, Cambridge Consultants, 1988.

[8] Hummel, R. A. and Landy, M. S. *A Statistical Viewpoint on the Theory of Evidence*. IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-10(2), pp. 319-325, 1988.

[9] Duff, M. J. B. *Review of the CLIP Image Processing System*. Proc. Natl. Computer Conf., pp. 1055-1060, 1978.

[10] Serra, J. *Mathematical Morphology and Image Analysis*. Academic Press, New York, 1982.

Defects	Lack of Root Fusion	Root Concavity	Gas Pore	Excess Penetration	Crack	Weld region
Location	Along medial axis of weld region	off-sided wrt weld region axis	anywhere within weld region	Anywhere witin weld region	longitudinal or transverse within weld	runs across image
Shape	very thin line like	longitudinally along weld axis; irregular	circular; spore-like;	light blob; irregular	fine, distinct; line-like;	horizontal rectangular region
Size	very thin	5-8 mm long 2-3mm wide	small; 0.3-0.5mm	2mm to half of the width of weld	0.2-0.3 mm	70mm long 10mm wide
Density	dark, darkness depend on depth of defect	darker than intensity of base metal	dark	light; density related to thickness of excess penetration	dark; transverse cracks are finer and not as dark as longitudinal cracks	light
Edge Definition	very sharp	sharp	sharp	badly defined;	sharp	not well-defined merge gradually with parent metal density
Acceptability	reject	depend on depth of concavity	depends on size of pores and distance between pores	depends on thickness of excess penetration	reject weld	
Notes	can be confused as a crack; can occur between interpass	severity of defect can be assessed from the size and density of defect region	cluster of pores is referred to as porosity	similar region which extends across entire width of weld is normal: weld capping	longitudinal cracks can be seen as multiple fine, distinct, and disjoint lines	

Figure 1 Image Characteristics of a sample of Weld Defects

Rules for infering object shape, location and orientation:

```
If      object_width/object_length > 0.7
and     object_area is NOT noise_like
                                                then object_shape is circular

If      object_width/object_length <= 0.7
and     object_area is NOT small
                                                then object_shape is elongated

If      object_area <= 3
and     object_location is near_image_boundary
                                                then object is noise_like

If      object is NOT noise_like
and     object_area < 10
                                                then object_size is small

If      object_skeleton_length > LONGthen skeleton_length is long

If      object.gradient < 30
and     object is NOT small
                                                then object is horizontal

If      abs_diff(object_centroid - weld_centroid)
        > SIMILAR
                                                then object is off_weld_axis
```

Figure 2. Example of Rules used in Iconic-Symbolic Conversion

```
Evidence (attribute value) => { hypothesis }:

Medial_axis_length (long) => {crack,RC,root,IQI,EP,LRF}
density wrt weld (different) => {crack,RC,root,pore,EP,LRF}
shape(elongated) => {crack,RC,root,IQI,LRF}
width(thick) => {RC,root,LRF}
size(medium) => {crack,RC,root,IQI,EP,LRF}
orientation(horizontal) => {crack,RC,root,EP,LRF}
location(near weld axis) => {cracl,root,IQI,pore,inclusion,LRF}
spore-like(false) => {crack,RC,root,IQI,pore,inclusion,EP,LRF}
dark feature(true) => {crack,RC,pore,LRF}

* DS ** -- singleton hypothesis [support, plausibility]

{crack} - [0.00, 0.349]
{RC} - [0.00, 0.189]
{root} - [0.00, 0.065]
{IQI} - [0.00, 0.004]
{pore} - [0.00, 0.003]
{inclusion} - [0.00,0.00]
{EF} - [0.00,0.002]
{LRF} - [0.49,1.00]
{} - [0.00,1.00]

Identification:

Defect candidate(s) with most support (0.49): {LRF}
Defect candidate(s) with most plausibility (1.00): {LRF}
Number of hypotheses generated: 19
Identification is LRF
```

Figure 3. System Output of the Identification