Attribute Impact for a Seismic Image Fusion System Based on Fuzzy Rules

Lionel VALET, Gilles MAURIS
LISTIC - ESIA
Université de Savoie, BP 806
74 016 Annecy Cedex
FRANCE
{valet,mauris}@univ-savoie.fr

Abstract – *This paper presents an attribute fusion method applied to 3D seismic block analysis. Attributes are extracted from the underground data and are aggregated by means of the fuzzy set theory which allows to code the linguistic attribute relation given by geophysicists under the form of IF-THEN rules. A cooperative system has been designed in order to involve the user in the fusion process. The main contribution of this paper concerns the elaboration of elucidative functionalities aimed at giving some explanations to the end-user on the obtained detection. An attribute impact based on the information theory is proposed in order to improve the decision interpretability.*

Keywords: Elucidative functionalities, cooperative fusion, information theory, knowledge representation, fuzzy systems.

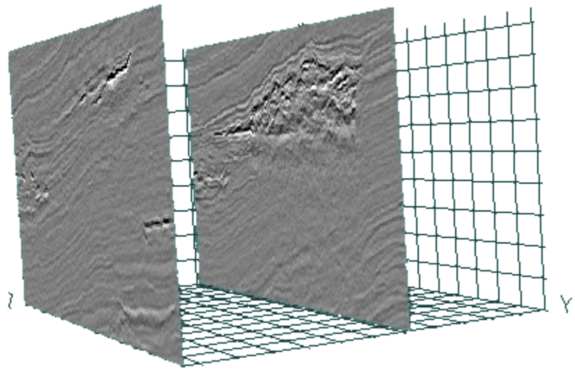


Fig. 1: A three dimensional seismic block.

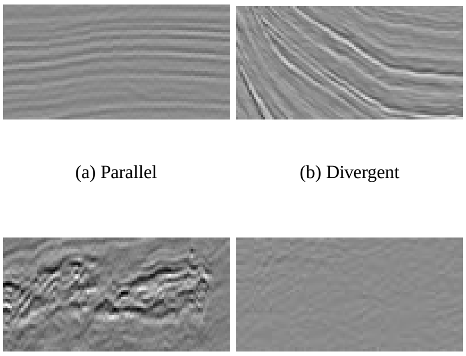
1 Problem presentation

The application presented in this paper is an industrial problem submitted by the Total oil company which deals with subsoil prospection. A step in the petroleum exploration chain is the analysis of three-dimensional images representing the underground strata. The data are acquired by the seismic reflection method which consists in generating sound waves by a surface shock. The waves are propagated underground and they are reflected by the different layers which compose the subsoil. With highly-sensitive geophones, the geophysicists record wave echoes and then, powerful computers generate tri-dimensional blocks representing the subsoil. Figure 1 shows an original block with only two vertical sections represented. This seismic block is a sample and corresponds to the following soil dimensions: about 2.3 km in depth, 2.6 km in length and 1.9 km in width.

The analysis of these data allows to determinate the subsoil organization. Expert interpretation consists in determining typical regions of the seismic image which are due to internal characteristics of the sound wave reflection used to constitute the block. Figure 2 shows some different types of seismic facies encountered.

The problem of seismic image segmentation is complex and a unique entity (attributes, measurements, . . .) to solve it is not available yet. Therefore, an approach based on information fusion [1] is used. It consists in merging attributes extracted from the 3D seismic block in order to solve the classification problem [2]. In such an application, there are no well-known mathematical relations between the attributes but interpreters of seismic images have an important knowledge on the sought-after regions. They are able to describe the regions using linguistic rules. For this reason, the aggregation is based on the fuzzy subset theory which is an adapted mathematical tool to code expert knowledge by means of IF-THEN rules. The main drawback of such a system is that many parameters must be set in order to be efficient. Thus, in order to propose a cooperative system easy to handle by interpreters, a graphic user interface has been developed [2]. It allows the user to benefit from the linguistic aspect of the fuzzy approach without any special training on the fuzzy subset theory.

In other respects, the interpretability of the decision represents a key point in this industrial application. The interpreters need to understand what occurs in the system [3] and they also need to have a great confidence in the obtained results. A way to help in the interpretation of the decision lies in a quantitative evaluation of the attribute impact on the final classification [2]. This notion, namely *elucidative fusion system*, was introduced by Dasarathy [3, 4] as a mean of understanding the relative contributions of the individual data source in the context of information fusion. Similar ideas have been previously developped in the artificial intelligence domain [5, 6] but with the aim to produce arguments of the reasoning drawn. Here the proposed approach concerns the information sources and is based on the comparison of the global information quantity (in the information theory sense) contained in the system with the



(a) Parallel (b) Divergent

(c) Chaotic (d) Dull

Fig. 2: Different kinds of facies encountered in seismic images.

local information characterizing each input.

The paper is organized in the following way: the next section presents the seismic attribute fusion based on fuzzy rules. The cooperative method used to adjust the membership functions is also described. Then section 3 focuses on the elucidative functionalities and section 4 is devoted to concluding remarks.

2 Seismic attribute fusion

The seismic attribute fusion used in this application has been previously presented in [2]. This approach is briefly recall in this section in order to introduce the proposed elucidative functionalities presented in the next section which is completely dedicated to this fusion approach.

2.1 Attribute presentation

Seismic facies are complex regions which cannot be detected using a single attribute. Interpreters are able to describe facies using two principal notions: the energy which represents the local contrast in the image and the isotropy which represents the local organisation. The energy and the isotropy are evaluated by two attributes extracted from the seismic block. They are obtained by a tri-dimensional calculation on seismic blocks. For each voxel, gradient vectors are computed in order to represent the local orientation. To measure locally the degree of organization, a principal component analysis of gradient vector fields yields three eigen vectors. They are sorted in decreasing order according to their respective eigen values: $\lambda_1 > \lambda_2 > \lambda_3$. Thus $\lambda_2 + \lambda_3$ is a measurement of isotropy and represents the notion of direction irregularities. $\sum \lambda_i$ represents the total energy and constitutes the second attribute.

2.2 Fuzzy fusion system

The fusion approach [1] is used in this application in order to realize a processing having the same behavior as experts (Fig. 3). The attributes are aggregated thanks to the fuzzy subset theory [7, 8]. This theory provides appropriate tools to code the interpreter reasoning by means of IF-THEN rules [9]. These rules are expressed in a linguistic way by interpreters. For example a rule which allows to detect chaotic zones is the following.

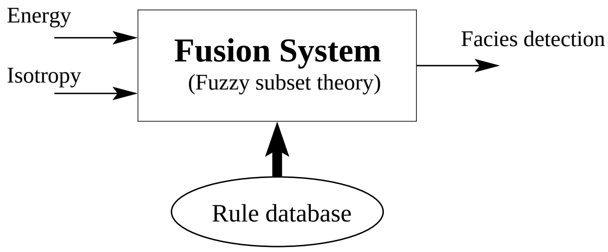


Fig. 3: Synopsis of the fusion system used to detect seismic facies.

IF isotropy IS high AND IF energy IS high
THEN voxel BELONGS TO chaotic region.

The fuzzy representation of this rule is made by a symbolic conjunctive view of the rules [10]. The rules are represented by a relation R such that $\mu_R(L_{energy}, L_{isotropy}, S)$ is equal to one if the rule is activated, to zero otherwise, e.g. $\mu_R(\text{high}, \text{high}, \text{chaotic}) = 1$.

A rule is composed of several words which describe each attribute. The larger the number of words is, the finer the description is. In this application, five words are used because interpreters work with five classes. The set of words which defines the variable isotropy is noted as follow:

$$\mathcal{L}_{isotropy} = \{\text{“very low”, “low”, “medium”, “high”,“very high”}\} \quad (1)$$

For each word, the associated membership function (its meaning) is defined according to interpreters' evaluations. Figure 4 shows the meaning of terms to the x_i attributes ($i = 1$ for the energy attribute and $i = 2$ for the isotropy attribute).

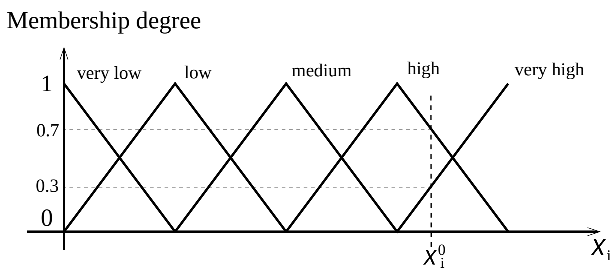


Fig. 4: Isotropy word meanings.

The linguistic description of a numeric value of an attribute is realized thanks to the word meanings. It consists, for a given input, in determining a membership value for each word according to its membership function. For example the numerical value x_i^0 in fig 4 for the isotropy attribute ($i = 2$) have the following Linguistic Description (LD):

$$LD(x_2^0) = 0.0/\text{very low} + 0.0/\text{low} + 0.0/\text{medium} + 0.7/\text{high} + 0.3/\text{very high} \quad (2)$$

$\mu_{LD(x_2^0)}(\text{very high}) = 0.3$ represents the membership degree of the value x of the isotropy attribute for the word "very high". Note that $\sum_{\forall k \in \mathcal{L}_{x_2}} \mu_{LD(x_2^0)}(L_k) = 1$ which characterizes the probabilistic behavior of interpreters.

Once the fuzzy linguistic descriptions of the inputs ($LD(x_i)$) are established, the next step consists of merging them to obtain an information regarding region belonging. The images of the input attributes by this relation R are obtained by means of the combination-projection principle [11] applied to the linguistic descriptions:

$$\mu(S) = \perp_{\forall L_{x_1}, L_{x_2} \in \mathcal{L}_{x_1}, \mathcal{L}_{x_2}} \top\Big(\top(\mu_{LD(x_1^0)}(L_{x_1}), \mu_{LD(x_2^0)}(L_{x_2})), \mu_R(L_{x_1}, L_{x_2}, S)\Big) \quad (3)$$

In this formula, the $\top$ norm is a combination operator and $\perp$ co-norm is a projection operator. There are a lot of $\top$ and $\perp$ operators [7]. It seems natural to choose them such that $\sum_{\forall S \in \mathcal{L}_{result}} \mu(S) = 1$ is satisfied in order to have the same property as for the inputs. In [10] a $\top$ norm and $\perp$ co-norm which allow to respect this constraint are studied and it leads to:

- $\top(a, b) = a * b$
- $\perp(a, b) = min(a + b, 1)$

The role of interpreters will consist in initializing the fusion system thanks to their experience by declaring the rules and by adjusting the word membership functions. A solution greatly valued is a graphic user interface easy to understand [2]. This interface is composed of a two-dimensional interactive table equipped with specific software functionalities. Rules are selected by means of a simple mouse click and membership functions are set in moving the horizontal and vertical lines of the table. Figure 5 illustrate the table construction.

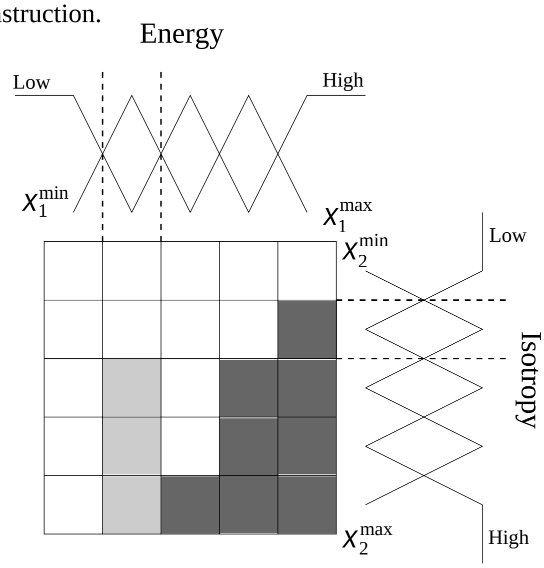


Fig. 5: The table rule construction principle.

2.3 Facies detection results

In this section, results obtained with the above mentioned fusion system are presented. Figure 6 is an original section from the seismic data block of figure 1. The result of classification given by interpreters is displayed in Figure 7. It should be noticed that it is a crisp result. In order to present the detection in a same way than interpreters expected, a defuzzyfication method based on the maximum membership degree has been applied to the computed output degree $\mu(S)$. The obtained detection is presented on the Figure 8. The chaotic facies are represented by black pixels and the dull facies by grey pixels. This result is coherent with the reference region of the figure 7. Some false alarms appear in the bottom of the image but they correspond to regions having the same property than the sought-after regions.

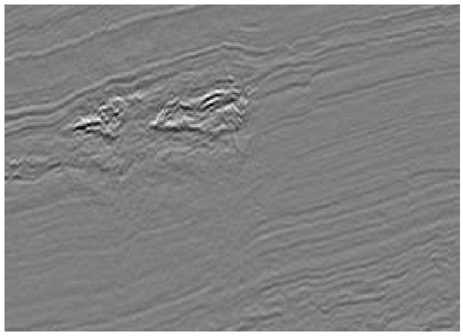


Fig. 6: One original section.

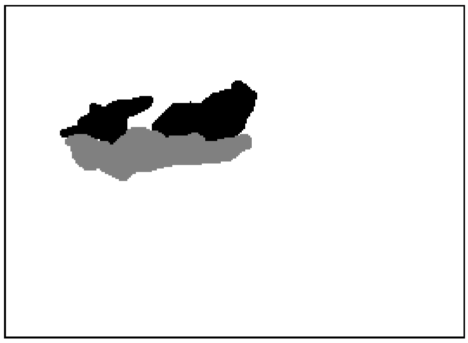


Fig. 7: The detection pointed by the interpreter.

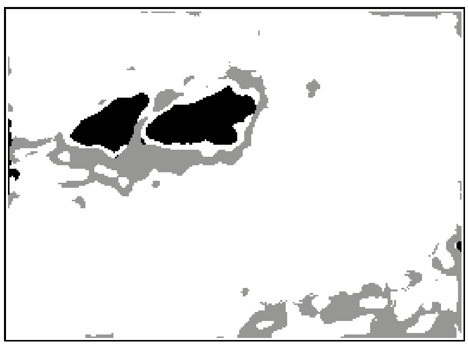


Fig. 8: The detection provided by the fusion system.

3 Elucidative functionalities

In such an industrial application, the need of interpretability of the obtained result is important. An analysis of the fuzzy fusion system is presented in order to have a better comprehension of the role of each attribute in the detection. The proposed approach is based on the evaluation of the attributes impact on the fusion result by the comparison of the information brought (in the information theory sense) by each attributes to the global information contained in the fuzzy rule database [3, 12].

3.1 Rule base information quantity

The first step consists in defining the global information quantity according to the rules given by expert. This information quantity is realized by integrating the global area of the rules for each output classes. A rule $\mu_R(A, B, C_j) = 1$ is represented by the intersection, noted $\mu_{C_j}(\vec{x})$, of two membership functions ($\mu_A(x_1)$ and $\mu_B(x_2)$) (see Fig. 5). The vector $\vec{x}$ represents the input attributes.

$$\vec{x} = \left| \begin{array}{l} x_1 = energy \\ x_2 = isotropy \end{array} \right. \tag{4}$$

The $\mu_{C_j}(\vec{x})$ intersection is realized by the product operator as presented in the previous section.

$$\mu_{C_j}(\vec{x}) = \mu_A(x_1) \times \mu_B(x_2) \tag{5}$$

In the case where several rules are used to define the output class C_j, $\mu_{C_j}(\vec{x})$ become the sum of all the sub-intersection corresponding to each rule. This 2D membership function $\mu_{C_j}(\vec{x})$ quantifies the membership degree of input $\vec{x}$ to the output class C_j. The quantity $V(C_j)$ in Eq. 6 represents the average membership degree of input patterns $\vec{x}$ to output class C_j over the studied domain $D \in [x_1^{min}, x_1^{max}] \times [x_2^{min}, x_2^{max}]$

$$V(C_j) = \frac{\int\limits_{\vec{x}\in D} \mu_{C_j}(\vec{x}) d\vec{x}}{\int\limits_{\vec{x}\in D} d\vec{x}} \tag{6}$$

The input membership functions are triangular so they are defined by three parameters presented in figure 9 for the word denoted A. Moreover, the partitions are strict which

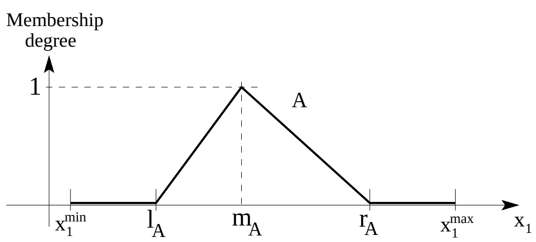


Fig. 9: Membership function representation.

allow to express the $V(C_j)$ in the following form:

$$V(C_j) = \frac{\sum\limits_{\forall A\in\mathcal{L}_{en}} \sum\limits_{\forall B\in\mathcal{L}_{iso}} (r_A - l_A)(r_B - l_B)\mu_R(A, B, C_j)}{4(x_1^{max} - x_1^{min})(x_2^{max} - x_2^{min})} \tag{7}$$

The product operator used for the aggregation allows to keep, in this 2D space, the 1D property $\sum\limits_{\forall A} \mu_A = 1$ for which it was selected:

$$\sum_j V(C_j) = 1 \tag{8}$$

Note that the relation R must be surjective for the satisfaction of this property, therefore a background class gathering all the non activated boxes (white boxes in Fig. 5) is defined.

The information quantity defined by Shannon [13] can thus be applied:

$$I(C) = -\sum_j V(C_j) log_2 V(C_j) \tag{9}$$

$I(C)$ is the global information quantity contains in the rule database. The global information quantity is maximum when input rules are equally spread to the output class. In this case, the entropy function applied to average value $V(C_j)$ is an interesting indicator. Indeed, an output class C_j having average value $V(C_j)$ equal to 1 means that all the rules of the input domain are devoted to this output class. In opposite, an average value $V(C_j)$ equal to zero indicate an output class that is never related with any rules in the input domain. The global information quantity obtained from the rule database of the Figure 5 is equal to 0.82. This indicator gives a first information concerning the input attribute choice according to the rules and the membership functions placement. The obtained value (0.82) is strong which means that the attributes (energy and isotropy) bring relevant information according to the representation of the phenomena. If a weak value is obtained at this fusion stage, changing or adding one or several input attributes has to be considered.

3.2 Word information quantity

The local information which characterizes each input is then evaluated also according to the rule database and to the membership functions. The input space $\vec{x} \in [x_1^{min}, x_1^{max}] \times [x_2^{min}, x_2^{max}]$ is divided into several subspaces according to the intersection between two consecutive membership functions as it is illustrated on the Figure 5. Each subspace is visualized by a square in the rule table and they are affected to a given output class by the interpreter.

For each word L_k of an attribute x_i (a raw or a column in the rule table), the average membership degree to each output class C_j is obtained with:

$$v(C_j|x_i = L_k) = \int\limits_{\vec{x}} \mu_{C_j|x_i=L_k}(\vec{x}) d\vec{x} \tag{10}$$

For example, for the word low of the energy attribute, the formula 10 become:

$$v(C_j|energy = \text{“}low\text{”}) = \sum_{\forall B \in \mathcal{L}\{Isotropy\}} \frac{3}{64}(r_{low} - l_{low})(r_B - l_B)\mu_R(low, B, C_j) \quad (11)$$

These values are normalized in order to evaluate their relative contribution in a probabilistic setting.

$$V(C_j|x_i = L_k) = \frac{v(C_j|x_i = L_k)}{\sum_j v(C_j|x_i = L_k)} \quad (12)$$

Then, the entropy function is applied to the obtained average membership degree in order to evaluate the local contribution of the word L_k for input x_i on the output set C:

$$I(C|x_i = L_k) = -\sum_j V(C_j|x_i = L_k) log_2 V(C_j|x_i = L_k) \quad (13)$$

A null value means that all the rules concerned by the world L_k correspond to the same output class. In opposite, a maximum value (which is equal to $log_2(Card(C))$) represents an equally spread of the rule output conclusions.

3.3 Attribute impact definition

Finally, the information brought by an attribute is defined by the average of the local contributions of each word describing the considered attribute:

$$I(C|x_i) = \frac{1}{Card(\mathcal{L}\{x_i\})} \sum_k I(C|x_i = L_k) \quad (14)$$

The attribute impact is built using the global information $I(C)$ (Eq. 9) and the local information $I(C|x_i)$ (Eq. 14). In practice, a set of word describing an input x_i for which all the output conclusions corresponds to the same sought-after class (without any influence of the other attribute, $I(C|x_i) = 0$), is considered to have a strong impact on the output. So, an expression of the input impact which follows this expected behavior is:

$$I_{x_i} = \frac{I(C) - I(C|x_i)}{I(C)} \quad (15)$$

A great value of the impact ($I_{x_i} = 1$) represents a great influence of the considered attribute in the detection, whereas a low value ($I_{x_i} = 0$) means that the input attribute has no influence on the results.

The Figure 10 presents a simple rule database. In this example, the output set is composed of two classes $C = \{C1, C2\}$. It is clear that the first attribute x_1 does not brought any information because all the rules are the same on each table raws. The global information computed for the domain $D \in [0, 255] \times [0, 255]$ and with equally spread membership functions gives $I(C) = 0.56$. The obtained input impacts are $I_{x_1} = 0$ and $I_{x_2} = 1$ which correspond to the expected impacts.

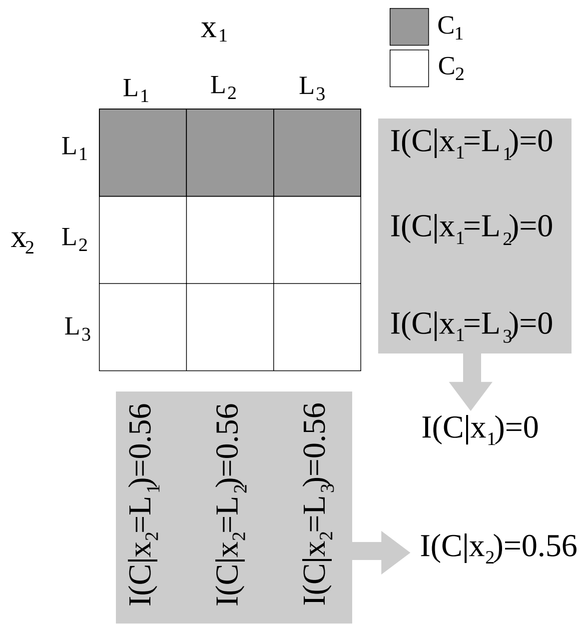


Fig. 10: attribute impact illustration

3.4 Application to seismic facies detection

The proposed input impact evaluation is implemented on the fuzzy fusion system dedicated to seismic facies detection presented in section 2. The results presented below are computed on the adjuted fusion system which has given the result of the figure 8.

The global information quantity $I(C)$ obtained from the rule database is equal to 0.82 which is a significant value. This quantity is computed in real time during the adjustment step which allows experts to follow the information quantity contained in the database when they modify the rules or the membership function placement. It helps in the fusion system parameter setting.

The word information quantities are given in the following table:

attribute	L_1	L_2	L_3	L_4	L_5
x_1	0.0	0.5	0.39	0.5	0.14
x_2	0.0	0.59	0.78	0.78	0.43

The word L_1 of the x_1 attribute (the energy attribute) brings no information on the sought-after regions since there are no rules which describe these regions (see Fig. 5). On the opposite the word L_3 of the x_2 attribute (the isotropy attribute) brings significant information since this word is involved in several rules.

The input impact evaluation, presented in the table below, shows that the energy has a more important impact ($I_{x_1} = 0.62$) than the isotropy attribute ($I_{x_1} = 0.37$).

Attribute	$I(C|x_i)$	I_{x_i}
x_1	0.30	0.62
x_2	0.51	0.37

Moreover, the result of the figure 8 is judged satisfactory by experts, which means that the fusion system is well adjusted. In this case, the impact values bring information about the importance of each input in the result. The two obtained impacts are significant which shows the interest of using this attribute pair in the detection of the sought-after regions. Moreover, the experts know that the result of the figure 8 is mainly based on the energy attribute thanks to its more important impact value. An immediate consequence is that this attribute has to be reliable. Finally, these indicators also brings a comparison base to the fusion of other attributes.

These impact indicators are computed in real time during the parameter adjustment step in order to help the end-user in this task. Indeed, the end-users are experienced experts and they have some knowledge on the mixing coefficients of various input attributes according to a priori information (the facies type, the country where the data come from, . . .). They can control the adjutment according to the attribute impact

4 Conclusions

The industrial application presented in this paper is a typical problem in which experts play an essential part in the decision process. Such experts have to justify the obtained results and the financial consequence are often important. For this reason, they do not have confidence in an automatic system which is a black box.

The proposed approach involves the interpreters in the detection by using their knowledge on the sought-after regions. The fuzzy subset theory is thus an appropriate tool to code interpreter knowledge under the form of IF-THEN rules. The rule database is then used to aggregate the input attributes in the same way as the interpreters themselves. In order to acquire the expert knowledge, a cooperative system is designed. It is based on an easy-to-handle graphic user interface. This interface allows non experts in the fuzzy subset theory to use the system by giving the rules of the sought-after regions and by adjusting the membership functions. The fusion is achieved in real time with the adjustment which allows to display the corresponding result immediately.

To help interpreters in the interpretation of the obtained results, an attribute impact based on the information theory is proposed. The information quantity of the fuzzy system is evaluated at several levels: the rule level, the word level and the attribute level. A combination of these different quantities is used to define the attribute impact, thus providing information about the reasons for the obtained classification. The obtained impacts bring some quantitative elements for justifying the obtained results. They also help the end-users in the fuzzy parameter adjustment step. The global information quantity and the attribute impacts are computed in real time during the adjustment step and they allow to know the effect of a rule modification or a membership placement immediately.

In this particular application where experts have an important knowledge of the sought-after regions, the number of rules is reasonable. Automatic construction of fuzzy model from a set of training examples is possible but it provides numerous fuzzy rules to offer sufficiently reliable performance for a low dimensional input space. The interpretability of such fuzzy models needs an automatic working description. Work is also under progress in order to use the quantities defined in this paper to select the attributes when the input dimensionality increases. Only attributes having a significant impact on the result can be kept for the fusion process.

Acknowledgements

The authors gratefully thank N. Keskes from the Total Company for providing the 3D seismic images and for his expert analysis.

References

[1] D. L. HALL. *Mathematical techniques in multisensor data fusion*. Artech House, 1992.

[2] L. VALET, G. MAURIS, P. BOLON, and N. KESKES. A fuzzy rule-based interactive fusion system for seismic data analysis. *Information fusion*, 4(2):123–133, June 2003.

[3] B. DASARATHY. Elucidative fusion systems - an exposition. *International journal on information fusion*, 1(1):5–15, 2000.

[4] R. SILIPO and M. BERTHOLD. Input features'impact on fuzzy decision processes. *IEEE Transactions on Systems, Man, and Cybernetic – part B: cybernetics*, 30(6):821–834, 2000.

[5] W. CLANCEY. The epistemology of a rule-based expert system – a framework for explanation. *Artificial Intelligence*, 20(3):215–251, 1983.

[6] L. AMGOUD and C. CAYROL. A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence*, 34:197–215, 2002.

[7] I. BLOCH. Information combination operators for data fusion: a comparative review with classification. *IEEE Transactions on Systems, Man, and Cybernetic – part A: systems and humans*, 26(1):52–67, 1996.

[8] R. HATHAWAY, J. BEZDEK, and W. PEDRYCZ. A parametric model for fusing heterogeneous fuzzy data. *IEEE Transactions on Fuzzy systems*, 4(3):270–281, 1996.

[9] D. DUBOIS, L. FOULLOY, S. GALICHET, and H. PRADE. Performing approximate reasoning with words ? In *Computing with words in Information/Intelligent systems 1*, pages 24–49. Springer verlag, 1999.

[10] G. MAURIS, E. BENOIT, and L. FOULLOY. The aggregation of complementary information via fuzzy sensors. *Journal of the International Measurement Confederation*, 17(4):235–249, 1996.

[11] L. ZADEH. Fuzzy sets. *Information and control*, (8):338–353, 1965.

[12] R. SILIPO and M. R. BERTHOLD. Input features'impact on fuzzy decision processes. *IEEE Trans. on Systems, Man and Cybernetics - Part B: Cybernetics*, 30(6):821–834, 2000.

[13] C. SHANNON. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 1948.