Identification of Choquet integral's parameters based on relative entropy and applied to classification of tomographic images

S. Jullien[1], G. Mauris[1], L. Valet[1], Ph. Bolon and S. Teyssier[2]
[1]Laboratoire d'Informatique, Systémes, Traitement de l'Information et de la Connaissance
{sylvie.jullien, lionel.valet, gilles.mauris, philippe.bolon}@univ-savoie.fr
[2]DST - Lab. de Recherche Matériaux, Rue Henri Tarze 38050 Grenoble
sylvie2.teyssier@schneider-electric.com

Abstract

A method based on relative entropy is proposed to identify Choquet integral parameters for classification purposes. While the entropy-based method initially developed by Kojadinovic [11] assigns a high importance to the attributes which allow to differentiate many classes, the relative entropy-based method proposed focuses on the researched classes. This is done in order to obtain the attribute confidence by the integration of expert knowledge concerning these classes. The values of the relative entropy for different subsets of attributes are considered as the values of Choquet capacity coefficients. The proposed identification method is integrated in an aiding system aiming at interpreting 3D-tomographic images of electrotechnical parts made of composite materials and manufactured by Schneider Electric. Four attributes extracted from the tomographic images are aggregated by the Choquet integral using the parameters identified by the proposed method. The result is a 3D cartography of the sought-after regions within the tomographic image. Quantitative assessments of classification highlight the relevance of the proposed approach.

Keywords: Identification, Choquet integral, relative entropy, classification.

1 Introduction

In classification and pattern recognition, relevant information, called features, and concerning classes of interest about an object is first extracted. Then, it is transformed into membership degrees according to the different classes and finally, it is combined (aggregation process) in order to improve decision-making concerning the class of the object.

In many applications, there are interactions between the features [15, 3, 16, 10] and these interactions bring complementarity or redundancy in the information fusion process. In order to take these interactions into account, the aggregation tool called Choquet Integral [9], relying on fuzzy measures, has given good results in many practical problems [3, 15, 4, 7].

The fuzzy measures used in the Choquet Integral enable to weight the importance of each subset of attributes. Therefore, in order to apply this powerful aggregation tool, identifying the fuzzy measures is required. However, it is a difficult problem [8] since it is equivalent to determining $2^N - 2$ parameters for N attributes of information.

In some practical applications, experts are able to provide some values of the fuzzy measure but it requires a deep knowledge of the information attributes. When the parameters cannot be provided by a priori knowledge, automatic methods have been proposed by several authors [5].

In the decision-aiding field, these methods generally rely on the concept of *preference* between attributes [17, 4, 14, 7]. In the classification field, the concept of *confidence* in the attributes is preferred in order to identify the importance of each subset of attributes [1]. A few methods have been developed in this view [3, 2] which this paper focuses on.

The recent identification method proposed by Kojadinovic [11] consists in assigning a weight to a subset of attributes according to its capability, or power, of discrimination between the different classes. This power of discrimination is determined by the entropy measure associated to the probability distribution of membership class degrees.

In supervised classification and pattern recognition, expert knowledge is available in the form of an annotated learning set. In this context, how is this prior information to be used in Kojadinovic's entropy-based identification method. In this paper, we propose a solution for this problem, based on *relative entropy*, that extends Kojadinovic's approach. The fuzzy measure is computed by the relative entropy between:

- the probability distribution provided by a set of attributes in referenced data annotated by experts as *the class associated to the Choquet Integral considered*,
- and the probability distribution provided by the same set of attributes in referenced data annotated by experts as *the complement to the class considered*.

The relative entropy (also named the Kullback-Leibler divergence) here quantifies the distance between a distribution associated to a class and a distribution associated to the other classes. Therefore, the identification method of Choquet Integral parameters uses, at the same time, positive and negative examples as in the powerful support vector machine (SVM) learning process well-known in machine learning [16]. In this approach, a different set of Choquet integral parameters is identified for each considered class. Finally, the classification decision simply consists in choosing the class that has the maximal degree. We also propose a thresholding process in order to create a class of rejects that is important in classification. The threshold value is called a degree of severity that is set by experts.

This learning method has been assessed on an industrial application in collaboration with Schneider Electric company where the goal is to detect regions of interest in 3D-tomographic images [10].

This paper is organised as follows: Section 2 describes the Choquet Integral, Section 3 presents the learning process of Choquet Integral parameters and finally Section 4 is dedicated to experiments.

2 Choquet capacities and Choquet Integral

Let $N = \{A_1, A_2, \dots A_i, \dots A_n\}$ denote the set of attributes. A Choquet capacity is a fuzzy measure that weights the importance of a subset of attributes $G = \{A_i, ..., A_j\} \subseteq S$ and is defined by:

$$\begin{array}{rcccl} \mu & : & 2^N & \rightarrow & [0,1] \\ & & G & \mapsto & \mu(G) \end{array} \quad (1)$$

The measure μ is a Choquet capacity if the following constraints are satisfied [9]:

- $\mu(\emptyset) = 0, \ \mu(N) = 1$
- $S \subseteq T \Rightarrow \mu(S) \leq \mu(T)$ (monotonicity)

The capacity is said to be:

- *additive* when $\mu(S \cup T) = \mu(S) + \mu(T)$, $\forall S, T \subseteq N \; such \; that \; S \cap T = \emptyset$ (probability measure),
- *super-additive* when $\mu(S \cup T) \geq \mu(S) + \mu(T), \forall S, T \subseteq N \; such \; that \; S \cap T = \emptyset$,
- *sub-additive* when $\mu(S \cup T) \leq \mu(S) + \mu(T)$, $\forall S, T \subseteq N \; such \; that \; S \cap T = \emptyset$.

In classification problems, the capacity is used in order to describe interactions between attributes [6]. One Choquet Integral is tuned for each class and each Choquet Integral aggregates the information provided by the attributes as follows [9]:

$$C(a_1, a_2, ..., a_n) = \sum_{i=1}^{n} (a_{(i)} - a_{(i-1)}) \cdot \mu(G_{(i)}) \quad (2)$$

where:

- $a_{(i)}$ are the values associated to class C_k provided by attribute $A_{(i)}$,

- $\mu(G_{(i)})$ is the importance of subset $G_{(i)} = \{A_{(i)}, ..., A_{(n)}\}$, and attributes are ordered according to the values $a_{(i)}$ associated to class C_k. The notation $_{(.)}$ indicates a permutation of indices according to the values provided by the attributes such as $a_{(i)} \leq a_{(i+1)} \leq ... \leq a_{(n)}$.

In some cases, the expression (2) is not adapted and too complex for practical problems. The approximated form called *2-additive Choquet Integral* has often been used and consists in considering a 2^{nd} order capacity [9] which only takes the weights of each attribute and the interaction between two attributes into account:

$$C(a_1, a_2, ..., a_n) = \sum_{i=1}^{n} \nu_i \cdot a_i - \frac{1}{2} \cdot \sum_{i=1}^{n} \sum_{j=i+1}^{n} I_{ij} \cdot \mid a_i - a_j \mid \quad (3)$$

where:

- ν_i is the weight A_i for the detection of class C_k,
- I_{ij} is the coefficient of interaction between both A_i and A_j for the detection of class C_k.

The parameters of the 2-additive Choquet Integral, i.e. ν_i and I_{ij}, can be obtained from the coefficient of importance $\mu(G_i)$ of a subset of attributes G_i by the following expressions [9]:

$$\nu_i = \sum_{T \subseteq N \setminus i} \frac{(n - |T| - 1)!|T|!}{n!} \times (\mu(T \cup \{i\}) - \mu(T)) \quad (4a)$$

$$I_{ij} = \sum_{T \subseteq N \setminus \{i,j\}} \frac{(n - |T| - 2)!|T|!}{(n-1)!} \times \Big(\mu(T \cup \{i,j\}) - \mu(T \cup \{i\}) - \mu(T \cup \{j\}) + \mu(T)\Big) \quad (4b)$$

3 Identification of Choquet Integral parameters

Several methods have been proposed for the identification of Choquet capacities in the context of decision-aiding [17, 4, 14, 11] by interpreting them as preferences for attributes that are aggregated. In the context of classification, the capacities are rather interpreted as degrees of confidence in attributes. Kojadinovic's entropy-based identification method [11] allows to solve classification problems. In the next paragraph, we describe this method and then we present our proposed *relative entropy*-based method.

3.1 Identification based on entropy

In Kojadinovic's entropy-based identification method [11], the confidence in each subset of attributes is interpreted as a discrimination power of the subset given a class. The larger number of classes a subset of attributes differentiates, the higher the confidence in this subset for the classification process. Therefore, the discrimination power is related to the diversity of membership degrees provided by a subset of attributes, which explains the use of entropy to quantify the confidence:

$$H^k(G_i) = \sum_{g_i} P(g_i) \cdot \log(P(g_i)) \quad (5)$$

where:

- $H^k(G_i)$ is the entropy of subset $G_i = \{A_i, ..., A_n\}$ describing class C_k,
- g_i gathers the vectors $\{a_i, ..., a_n\}$ of membership degrees to class C_k provided by the subset G_i,
- $P(g_i)$ is the distribution of joint probabilities associated to g_i,
- $H^k(N)$ is the entropy of subsets $\{A_1, ..., A_n\}$.

Moreover, when normalized as follows:

$$\mu^k(G_i) = \frac{H^k(G_i)}{H^k(N)} \quad (6)$$

the entropy measure satisfies the constraints of Choquet capacities. One can remark that there are as many Choquet capacities μ^k as the number of classes C_k.

This method of identification is based on the whole set of membership degrees without taking prior information concerning classes into account. Considering that the confidence of an attribute is related to the maximum number of classes that this attribute can differentiate, is not relevant when the number of researched classes is limited. Moreover, integrating expert knowledge on some membership degrees where the class is precisely known seems to be an interesting extension that could improve results. For that, we propose a relative entropy-based approach.

3.2 Relative entropy

The relative entropy takes the distributions $P(g_i^k)$ and $P(g_i^{\overline{k}})$ into account. They are computed from both data referenced as the class C_k and data referenced as classes different from C_k and denoted:

$$C_{\overline{k}} = \bigcup_{j \neq k} C_j \tag{7}$$

The relative entropy is therefore:

$$H_R^k(G_i) = \sum_{g_i^k} P(g_i^k) \cdot \log\left(\frac{P(g_i^k)}{P(g_i^{\overline{k}})}\right) \tag{8}$$

where:

- $P(g_i^k)$ is the distribution of probabilities of subset g_i^k of membership degrees to class C_k provided by the subset of attributes G_i and corresponding to class C_k,
- $P(g_i^{\overline{k}})$ is the distribution of probabilities of subsets $g_i^{\overline{k}}$ of membership degrees to class C_k provided by the subset of attributes G_i and corresponding to the other classes $C_{\overline{k}}$.

This approach strongly depends on prior knowledge brought by experts. In the industrial application considered in the experiment section, this information is provided by expert pointers under the form of graphical pointing.

3.3 From relative entropy to Choquet capacities

The relative entropy has to satisfy the conditions presented in Section 2 in order to be interpreted as a Choquet capacity. For that, the relative entropy $H_R^k(G_i)$ for a subset G_i is normalized as in Kojadinovic's method by the relative entropy of the whole set of attributes $H_R^k(N)$:

$$\mu^k(G_i) = \frac{H_R^k(G_i)}{H_R^k(N)} \tag{9}$$

Moreover, the relative entropy is null when the set G_i is empty but also when both distributions $P(g_i^k)$ and $P(g_i^{\overline{k}})$ are identical. Therefore, an attribute that provides the same membership degrees for a researched class C_k and for the other classes $C_{\overline{k}}$ is assigned a low importance value since it cannot distinguish class C_k from the others.

The relative entropy also has to satisfy the monotonicity constraint (Section 2). Given two attributes A_i and A_j, the relative entropy has to satisfy the following equations:

$$\mu(\{A_i, A_j\}) \geq \mu(\{A_i\}) \tag{10a}$$

$$\mu(\{A_i, A_j\}) \geq \mu(\{A_j\}) \tag{10b}$$

In order to check these constraints, the relative entropy can be rewritten as [12, 13]:

$$\begin{aligned} H_R^k(\{A_i, A_j\}) = \\ H_R(\{A_i\}) + H_R^k(\{A_j \mid A_i\}) \end{aligned} \tag{11}$$

and it is always positive [12] therefore it has a monotonic behaviour:

$$H_R^k(\{A_i, A_j\}) \geq H_R^k(\{A_i\}) \tag{12a}$$

$$H_R^k(\{A_i, A_j\}) \geq H_R^k(\{A_j\}) \tag{12b}$$

This reasoning can easily be extended to larger subsets of attributes. Therefore, the normalized relative entropy satisfies all the constraints in order to be interpreted as a Choquet capacity.

3.4 Modelling positive and negative interactions

When attributes A_i and A_j, that provide distributions $P(g_i^k)$ and $P(g_i^{\overline{k}})$, are independent, the relative entropy has an additive behaviour [13]:

$$H_R^k(\{A_i, A_j\}) = H_R^k(\{A_i\}) + H_R^k(\{A_j\}) \tag{13}$$

When attributes A_i and A_j are interacting with each other, the relative entropy can be expressed by:

$$\begin{aligned} H_R^k(\{A_i, A_j\}) = H_R^k(\{A_i\}) + H_R^k(\{A_j\}) \\ + \Big(H_R^k(\{A_j|A_i\}) - H_R^k(\{A_j\})\Big) \end{aligned} \tag{14}$$

where the last term can be negative or positive according to attributes A_i and A_j implying that the identified Choquet capacities can be super-additive or sub-additive. Therefore the proposed method allows to model and identify both positive and negative interactions whereas Kojadinovic's approach can only identify negative interactions [11].

3.5 Distributions support

In order to use the relative entropy, the support of the distribution $P(g_i^k)$ must be included in the support of the distribution $P(g_i^{\overline{k}})$. In the opposite case, the relative entropy diverges towards infinity. In order to satisfy this constraint, we propose to use the Skew divergence defined as:

$$D_S(G_i) = \sum_{g_i^k} P(g_i^k) \times \log\left(\frac{P(g_i^k)}{\alpha \cdot P(g_i^k) + (1-\alpha) \cdot P(g_i^{\overline{k}})}\right) \quad (15)$$

This Skew divergence consists in adding a low percentage α (e.g. 0.1) of the distribution $P(g_i^k)$ to distribution $P(g_i^{\overline{k}})$. This process allows to check the constraint concerning the distribution support. Since the Skew divergence is a particular type of relative entropy, it can be interpreted as a Choquet capacity.

4 Tomographic images interpretation

An information fusion system has been developed in order to give a hand to the experts of Schneider Electric company for electrotechnical part analysis based on a non-destructive process called 3D-tomography. The analysis of fiber organization within the parts enables experts to assess the quality of these parts. An example of tomography is illustrated in figure 1 where the experts have pointed some reference regions corresponding to the researched classes:

- oriented regions denoted C_1 where the fibers have similar orientations,
- disordered regions denoted C_2 where the fibers are mixed together,
- lack of reinforcement regions denoted C_3 that contain a few quantity of fibers.

Therefore, prior knowledge provided by experts is available for the classification.

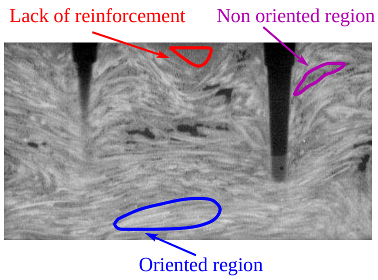


Figure 1: Pointed (referenced) regions given by experts.

In order to detect the sought-after regions, four attributes have been extracted from the original images (c.f. figure 1):

- Attributes A_1, A_2, A_3 provide information on fiber orientation using gradients of intensities by three different measures
- Attribute A_4 provides information on fiber homogeneity based on texture analysis using co-occurence matrix.

This information is not commensurate therefore its interpretation by experts is difficult. In order to solve this problem, the measures provided by the attributes have been transformed into *membership maps* expressed in the same scale ($[0,1]$). One map is computed for each sought-after region by a mechanism of comparison between pointed regions. The proposed mechanism is independent from the identification method and has been thoroughly described in [10]. These maps (see examples in figure 2) represent relevant information concerning the regions and present complementarity and redundancy that can be exploited by Choquet Integrals.

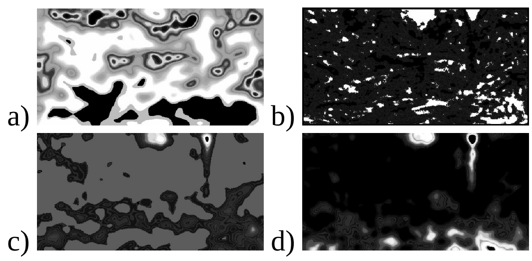


Figure 2: Membership maps of lack of reinforcement regions obtained from the attributes A_1, A_2, A_3 and A_4.

The membership degrees, such as the ones illustrated in figure 2, are aggregated using a 2-additive Choquet Integral (eq. 3). The reason for which we do not use the expression (2) is that the 2-additive form is more easily interpreted by experts since the interaction coefficients only concern two attributes at a time. The tables 1 and 2 summarize the parameters identified by the methods based respectively on the entropy and the relative entropy. The method based on the entropy uses the whole map, unlike the method based on the relative entropy which uses the regions pointed by experts. The probability distributions computed respectively on the lack of reinforcement regions and on the other regions and provided by attributes A_1 and A_4, are presented on the figure 3(a) and (b). The distributions computed on the attributes A_4 are rather similar. This attribute does not really allow to distinguish the oriented regions from the other sought-after regions. Unlike the attribute A_4, the distributions computed on the attribute A_1 are really different. This attribute allows to dissociate the oriented regions from the other regions.

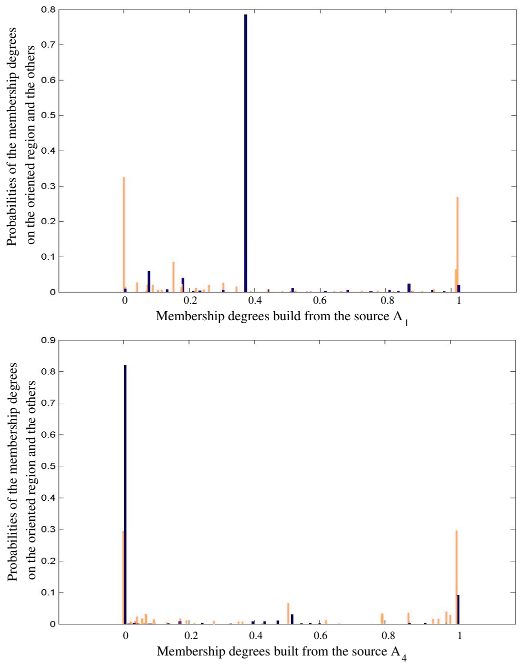
a)
b)


Figure 3: Probability distributions of membership degrees on the oriented region (clear color) and on the other regions (dark color) provided by two attributes A_1 and A_4.

The Skew divergence (eq. 15) has been applied to the probability distributions where $\alpha = 0.1$. The Choquet capacities identified have been converted into weights and interaction coefficients by the equations (eq. 4a) and (eq. 4b) to allow interpretation by the experts.

The parameters identified by the method based on the entropy leads to a quasi-weighted average (interactions are negligible in the table 2). The proposed method based on the relative entropy allows to determine the key attributes in the differentiation of the sought-after regions. The identification method based on the relative entropy allows to obtain more important interaction coefficients. The attributes A_1, A_2 and A_3 which characterize the orientation of intensity gradients are more important than the attribute A_4 which characterizes the texture homogeneity. That is correct because the orientation is the main characteristic of the oriented regions. Moreover, the identification method based on the relative entropy provides negative and positive interaction coefficients.

	Weights					
	C_1		C_2		C_3	
	H	H_R	H	H_R	H	H_R
A_1	0.05	0.49	0.12	0.48	0.17	0.14
A_2	0.42	0.29	0.18	0.39	0.56	0.25
A_3	0.10	0.15	0.14	0.05	0.17	0.34
A_4	0.43	0.07	0.56	0.08	0.10	0.27

Table 1: Weights identified by the method based on the entropy $H(G_i)$ and the method based on the relative entropy $H_R(G_i)$ for all the attributes A_1, A_2, A_3 et A_4.

	Interaction coefficients					
	C_1		C_2		C_3	
	H	H_R	H	H_R	H	H_R
A_1, A_2	-0.04	-0.35	-0.05	-0.61	-0.07	0.01
A_1, A_3	-0.01	-0.14	-0.01	-0.03	-0.04	-0.13
A_1, A_4	-0.01	-0.09	-0.02	-0.10	-0.01	-0.13
A_2, A_3	-0.04	-0.15	-0.01	-0.05	-0.07	-0.33
A_2, A_4	-0.05	-0.02	-0.03	-0.10	-0.02	-0.28
A_3, A_4	-0.03	-0.03	-0.03	-0.04	-0.02	-0.38

Table 2: Interaction coefficients identified by the method based on the entropy $H(G_i)$ and the method based on the relative entropy $H_R(G_i)$ for all the attributes A_1, A_2, A_3 et A_4.

The Choquet integral provides one membership map per class. Then, the cartography building is based on a thresholding. For example, the oriented regions on the cartography contain all the voxels which have a membership degree on the oriented region greater than the others and greater than a threshold named severity degree. This process is applied on each sought-after region. When all the membership degrees are lower than the threshold, the voxels are included in rejected regions. This threshold is set by the experts according to the severity level of the part study. As it is presented in the figure 4, the cartography contains white regions (lack of reinforcement), clear grey regions (non-oriented regions), dark grey regions (oriented regions) and black regions (rejected regions).

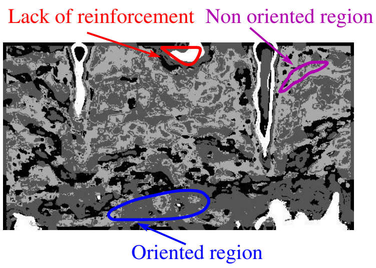


Figure 4: The pointed regions are shown on the cartography obtained from the entropy.

The cartographies of sought-after regions (c.f. figures 4 and 5) have been evaluated with the confusion matrix. The pointed regions used in the evaluation are different from the pointed regions used in the identification. Four performance measures have been extracted from the confusion matrix [18]: the classification rate, the recall, the precision and the F-measure.

Figure 5: The part cartography obtained from the relative entropy.

The classification rate evaluates the rate of voxels correctly classified out of all the voxels of the evaluation set. The recall quantifies the rate of voxels correctly classified out of all the voxels classified in a sought-after region. The precision is the rate of voxels correctly classified out off all the voxels in a sought-after region describe on the evaluation set. The F-measure summarizes the two last evaluations.

(a) The evaluation of the cartography obtained from the entropy.

	Recall	Precision	F-measure
C_1	80%	93%	86%
C_2	$\mathbf{13\%}$	$\mathbf{61\%}$	$\mathbf{21\%}$
C_3	99%	$\mathbf{37\%}$	$\mathbf{53\%}$

(b) The evaluation of the cartography obtained from the relative entropy.

	Recall	Precision	F-measure
C_1	80%	92%	85%
C_2	56%	$\mathbf{76\%}$	$\mathbf{64\%}$
C_3	99%	55%	$\mathbf{71\%}$

Table 3: The recall, precision and F-measure of two different cartographies.

The system composed of the identification method based on the entropy does not correctly detect all the regions pointed by experts (c.f. figure 4). Its classification rate on the pointed regions for evaluation is 50.8%. The identification method proposed in this paper, integrates expert knowledge to improve the cartography quality. The classification rate in the evaluation regions is improved to reach 93.8%. The method proposed allows to improve the classification rate by 43%. The tables 3a and 3b summarize the evaluations of the cartographies obtained from the two identification methods presented. The F-measures (computed with $\beta = 1$) show the improvement provided by the proposed identification method. The F-measures have increased respectively by 43% and 18% for the non oriented regions and the lack of reinforcement regions.

5 Conclusion

This paper presents an extension of the method proposed by Kojadinovic to identify the Choquet integral parameters. This method is applied specifically to the data classification field. The importance coefficients identified are interpreted as the confidence placed on each attribute. The proposed relative entropy allows to quantify the confidence according to the considered classes. Moreover, the interaction coefficients deduced from the Choquet capacities can be positive or negative according to the attributes and the sought-after regions. The identification method proposed is integrated into an aiding system of tomographic image interpretation. This system fuses attributes extracted from the tomographic images with the Choquet integral. The proposed method increases the quality of the cartography compared with the results obtained with an identification method based on the entropy. It will be interesting to analyse the sensitiveness of the identification method according to the regions pointed by experts and the building method of membership maps.

Acknowledgement

This study is a part of the FUSEXP French program supported by the “Région Rhône Alpes”. This program includes Schneider Electric Group, the “Cité de l’image en mouvement” and the LISTIC from the University of Savoie.

References

[1] J. Bezdek, J. Keller, R. Krisnapuram, N. Pal. Fuzzy models and algorithms for pattern recognition and image processing. *Kluwer Academic Publishers*, 1999.

[2] J.H. Chiang, Aggregating membership values by a Choquet-fuzzy-integral based operator. *Fuzzy Sets and Systems,* 114 : 367-375, 2000

[3] P.D. Gader, M.A. Mohamed, J.M. Keller. Fusion of handwritten word classifiers *Pattern Recognition Letters*, 17 : 577–584, 1996.

[4] M. Grabisch. A new algorithm for identifying fuzzy measures and its application to pattern recognition. *IEEE International Conference on Fuzzy Systems*, 1 : 145–150, 1995.

[5] M. Grabisch, I. Kojadinovic, P. Meyer. A review of methods for capacity identification in Choquet integral based multi-attribute utility theory Applications of the Kappalab R package *European Journal of Operational Research*, 189 : 766-785, 2008.

[6] M. Grabisch, M. Sugeno. Multi-attribute classification using fuzzy integral. *IEEE International Conference on Fuzzy Systems,* 1992, 47–54.

[7] M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89 : 445–456, 1996.

[8] M. Grabisch, I. Kojadinovic, P. Meyer. Using the Kappalab package for Choquet integral based multi-attribute utility theory. *Information Processing and Management of Uncertainty (IPMU)*, 2006.

[9] M. Grabisch, k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92 : 167–189, 1997.

[10] S. Jullien, L. Valet, G. Mauris, Ph. Bolon, S. Teyssier. Système d’aide a l’évaluation de la qualité de pièces en materiau composite basé sur l’integrale de Choquet. *Rencontres Francophones sur la Logique Floue et ses Applications(LFA)*, 2006.

[11] I. Kojadinovic. Estimation of the weights of interacting criteria from the set of profiles by means of information-theoretic functionals. *European Journal of Operational Research*, 155 : 741–751, 2004.

[12] S. Kullback, R.A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1) : 79–86, 1951.

[13] H. Le-Borgne, A. Guerin-Dugue, A. Antoniadis. Representation of images for classification with independent features. *Pattern Recognition Letters*, 25(2) : 141–154, 2004.

[14] J.L. Marichal, P. Meyer, M. Roubens. Sorting multi-attribute alternatives: The TOMASO method. *Computers and Operations Research*, 32 : 861-877, 2005.

[15] L. Mikenina, H.J. Zimmermann. Improved feature selection and classification by the 2-additive fuzzy measure. *Fuzzy Sets and Systems* 107 : 197–218, 1999.

[16] V. Mitra, C.J. Wang, S. Banerjee. Text classification: A least square support vector machine approach. *Applied Soft Computing*, 7(3) : 908–914, 2007.

[17] T. Mori, T. Murofushi. An analysis of evaluation model using fuzzy measure and the Choquet integral. *5rh Fuzzy Systems Symposium*, 1989 (in japanese).

[18] M. Sokolova. Assessing invariance properties of evaluation measures. *Proceedings of the 19th Neural Information Processing Systems Conference*, (NIPS 2006). (2006)