Proceedings of the 2007 IEEE Symposium on Computational
Intelligence in Multicriteria Decision Making (MCDM 2007)

A Spatial Classification Model for Multicriteria Analysis

Ana Del Amo *Member, IEEE*, Luis Garmendia, Daniel Gómez and Javier Montero

***Abstract*— This paper stresses that standard multicriteria aggregation procedures either do not assume any structure in data or this structure is in fact assumed linear. Nevertheless, many decision making problems are based upon a family of data with a well defined spatial structure, which is simply not taken into account. Hence, such aggregation procedures may be misleading. Therefore, we propose an alternative model where the aggregation of criteria assumes a certain structure, according to remote sensing data.**

I. Introduction

In the emerging area of sensor-based systems, an existing significant challenge is the development of scalable methods capable of extracting useful information from the data that the sensors collect. The extraction of useful information from different sources can be modeled as a Multicriteria decision making problem. Aggregation operators represent a standard mathematical tool in Multicriteria decision making. By means of an appropriate aggregation operator, information collected from different sources is amalgamated into a simpler descriptive index (see, e.g., [3]). As already pointed out in [15], a number of assumptions in aggregation operators are artificially introduced, apparently justified from a usefulness argument. For example, many researchers define aggregation in terms of a unique binary operation rule, which is sequentially applied under the associativity argument. This approach refers to a family of data that can be represented in the real line (according to time, for example). But in many decision making problems the data collected comes from a surface, real three-dimensional data or in some cases higher dimensions. In the case of a sensor based system the data could have been collected from a LADAR, SAR and satellite, all from the same area for example. This is the case when we analyze remotely sensed images, where pixels define a specific structure due to contiguity (see, e.g., [12], [19]).

When analyzing a pixel from a remotely sensed image, for example, we describe the pixel as a mixed of different degrees of membership to every main *concept* (forest or wetland, for example). Surrounding pixels will also be described by their degrees of membership to each one of the same main concepts. Therefore, it makes sense to conceive that the true degree of membership for each one of the pixels in the image as some kind of average obtained from each pixel and its surrounding neighbors. This kind of average will produce a natural smoothing, which in some context (depending on the required precision) will provide a more accurate description of the pixel than the one obtained from the single observed pixel, at least for decision making purposes. For example, a forest containing a small country house should still be defined as a forest. The country house will be perceived as an outlier pixel or as a set of outlier pixels. In this case, we may not need the information about the existence of these outlier pixels or we might decide not to take it into account in our decision process. Notice that we are suggesting a data reconstruction technique that may not be associated to any standard smoothing technique, where generally the observed value of a pixel is taken into account. We proposed to substitute the observed value by the aggregated value obtained from the set of surrounding pixels, in order to produce a simpler description of the image, easier to be understood and to be managed by the decision maker (see [14]).

The paper is organized as follows: the next section presents a short criticism of standard assumptions in aggregation operators. We show some of the consequences of these assumptions, specifically those related to the underlying structure of data being assumed. In section III, we establish the connection between some multicriteria models and the aggregative classification presented in [7]. We also proposed an alternative model that takes into account the spatial structure in which the data in a standard remotely sensed image is arranged. Finally, the relevance of this approach is discussed in section IV, where we take into account standard approaches in remotely sensed images.

II. Standard approaches to aggregation procedures

The aggregation of information is a key issue in a number of problems. The decision maker needs to reduce the complexity of the problems in which the information is obtained from multiple sources.

The formal approaches to this problem found in the literature are based on binary connectives, mainly t-norms, t-conorms and uninorms (see, e.g., [3], [32]). These families of binary connectives are conceived (see also [11], [20], [30]) as sets of mappings

$$\odot : [0,1] \times [0,1] \rightarrow [0,1]$$

verifying standard assumptions that include

1. Commutativity:

$$\odot(a,b) = \odot(b,a), \quad \forall a,b \in [0,1]$$

Ana Del Amo is in the Mission Management Department, Smiths Aerospace, Grand Rapids, Michigan 49512-1991, USA (phone: +1 616 2417982; email: ana.del.amo@smiths-aerospace.com).

Luis Garmendia is with Faculty of Informatics, Complutense University, Madrid 28040, Spain (phone: +34 91 394 7642; e-mail: lgarmend@fdi.ucm.es).

Daniel Gómez is with Faculty of Statistics, Complutense University, Madrid 28040, Spain (phone: +34 91 394 3994; email: dagomez@estad.ucm.es).

Javier Montero is with Faculty of Mathematics, Complutense University, Madrid 28040, Spain (phone: +34 91 394 4522; email: javier_montero@mat.ucm.es).

1-4244-0702-8/07/$20.00 ©2007 IEEE

2. Monotonicity:

$$a_1 \leq a_2, b_1 \leq b_2 \Rightarrow \odot(a_1, b_1) \leq \odot(a_2, b_2)$$

3. Associativity:

$$\odot(\odot(a,b),c)) = \odot(a, \odot(b,c)), \quad \forall a,b,c \in [0,1]$$

In addition, two *extreme* boundary conditions are also assumed:

4α. $\odot(1,1) = 1$
4β. $\odot(0,0) = 0$

Some other extra assumptions may be imposed in order to assure certain properties, perhaps desirable from a practical point of view:

5. Continuity: given a sequence $\{(a_n, b_n)\}_{n=1}^{\infty} \subset [0,1]$ such that $\lim_{n\to\infty}(a_n, b_n) = (a,b)$, then

$$\lim_{n\to\infty} \odot(a_n, b_n) = \odot(a,b)$$

6. Idempotency:

$$\odot(a,a) = a, \quad \forall a \in [0,1]$$

Moreover, particular conditions are associated to certain classical mathematical properties. For example, conjunction and disjunction are respectively fixed by the following exclusive boundary conditions:

7t. For t-norms:

$$\odot(a,1) = \odot(1,a) = 1, \forall a \in [0,1]$$

7c. For t-conorms:

$$\odot(a,0) = \odot(0,a) = 0, \forall a \in [0,1]$$

But as pointed out in [15], the majority of the above conditions seem to be accepted without discussion in most of the cases. For example:

- Commutativity refers to the irrelevancy of data ordering in the aggregation process, in such a way that the results of the aggregation will be the same independently of the permutation (see [21]). It is indeed a standard *mathematical* property, but it implies a severe restriction in some cases against the real description of the problem. Commutativity can be properly assumed only when our experiment has been designed in order to fulfill commutativity.
- Monotonicity, in principle, looks like an obvious property (if the degrees of truth increase in each part, the global degree of truth should increase, or at least, never decrease). But this property deserves much more attention, it can not be freely assumed without a robust supportive background. In fact, it is well known in Reliability Theory (see [1]) that a similar condition does not necessarily hold when dealing with physical systems subject to failure. Monotonicity indeed is a standard assumption in Reliability Theory, even in a non binary context (see, e.g., [2], [28]), but it should not be accepted without remarking that some interesting systems are being discarded (see [15] for an in dept discussion).
- In the case of associativity, we must agree with [20] that the very first definition of triangular norm, as proposed by Menger [22], did not required associativity. The associativity condition, was included later [30]. The problem is that, in order to be useful in practice, we need binary connectives, but we also need to be able to aggregate any arbitrary number of arguments. Associativity takes care of this issue, providing the tool to be able to extend each binary operator in a unique way, just by induction [20], independently of the direction of the calculus (left to right or right to left). Therefore, associativity can be envisioned as a necessary restriction, only when we impose the restriction over the aggregation process to be based on a unique binary operation. Such an assumption is not obvious in the crisp case, and it is very difficult to accept in a more general context.
  Of course, aggregation should take into account which pieces of information have already been aggregated. Montero, for example, in [23] and [26] includes the size of the aggregated information in order to avoid Fung and Fu restrictive result [13]: if we assume that the aggregation process must be based upon a unique binary operator, standard restrictions will drastically reduce the set of available binary operators.
  The aggregation processes should not only be based on binary operations but could *evolve* in time or could remain constant along the complete aggregation process. Defining an operational aggregation rule can be alternatively assured allowing a successive reckoning of binary operators, by means of recursiveness, as proposed in [4] and [7].
- For extreme boundary conditions we can apply analogous arguments to those above, relative to monotonicity: it may even happen that the concept of *best* and *worst* have to be revised in order to meet extreme boundary conditions.
- Continuity is sometimes considered as a very strong mathematical condition, and in fact there are a number of key results that can be obtained assuming some of the weaker forms of continuity (see, e.g., [20]). From a practical point of view, we should stress the weakness of the standard continuity. If the evaluation space is, in fact, the whole unit interval, a robust estimation associated procedure requires a stronger restriction. We do desire our binary connective to be *smooth*, but it can be derivable and still show very high slopes, in such a way that a *small* input measurement error can still produce a big change in the output. This situation is not avoided just assuming that our function is infinitely derivable. In practice, we should be imposing certain smoothness restriction (the absolute value of the first derivative must not be too high), and such a smoothness restriction most likely will depend on our precision measurement level.

Some smoothness restriction becomes important in *Soft Computing* systems.

- Idempotency is certainly violated in some contexts with some kind of *attraction* behavior: a high degree, if repeated, may suggest in many contexts a higher aggregated degree, and a low degree, if repeated, may suggest a very low aggregated degree. Distributive logics verify the absortion law, which imply idempotency. Note, for example, that in case of the conjunction and disjunction continuous connectives, only the operators of Zadeh's logic (maximum and minimum) hold idempotency, and so is the only logic using continuous operators that can be distributive. Logics using archimedean t-norms are not idempotent, and they do not hold the absortion law, and they do not hold the distributive property.
- Exclusive boundary conditions are introduced in order to give the aggregation a particular meaning. But again we should not be accepting that there are only two possible roles for binary aggregation, as in the crisp context. Alternative boundary conditions may lead to different aggregation operators, as binary weighted means or Yager's binary OWA operators [31], [32].

Moreover, we should never expect that a complex situation can be summarized by a unique index. Not even in the case that the multiple sources of information assume a common measurement unit (which may not be assured), and of course, we can not force the mix of data items that are absolutely different in nature.

Nevertheless, in order to try to solve some of the issues that we have presented above, we need to utilize recursiveness. The key idea of the concept of recursiveness is that an aggregation rule, in order to be operational, should be based upon an iterative application of binary operators, but not necessarily the same binary operator. Therefore, the data is assumed to be aggregated one by one, implying perhaps a previous re-arrangement or reordering of the data (see [4] for more details).

Definition 2.1: Let us denote a permutation

$$\pi_n(a_1, a_2, \ldots, a_n) = (a_{\pi_n(1)}, a_{\pi_n(2)}, \ldots, a_{\pi_n(n)})$$

An *ordering rule* π is a *consistent* family of permutations $\{\pi_n\}_{n>1}$ such that for any possible finite collections of numbers, each extra item a_{n+1} is allocated keeping previous relative positions of items, i.e.,

$$\pi_{n+1}(a_1, a_2, \ldots, a_n, a_{n+1}) = (a_{\pi_n(1)}, \ldots, a_{\pi_n(j-1)}, a_{\pi_{n+1}(j)}, a_{\pi_n(j)} \ldots, a_{\pi_n(n)})$$

for some $j \in \{1, \ldots, n+1\}$.

In other words, once the relative position of two elements is fixed by means of a permutation π_n, no other permutation π_m, $m > n$, will modify the relative position.

The following definition was proposed in [4].

Definition 2.2: A left-recursive connective rule is a family of connective operators

$$\{\phi_n : [0,1]^n \to [0,1]\}_{n>1}$$

such that there exists a sequence of binary operators

$$\{L_n : [0,1]^2 \to [0,1]\}_{n>1}$$

verifying

$$\phi_2(a_1, a_2) = L_2(a_{\pi(1)}, a_{\pi(2)})$$

and

$$\phi_n(a_1, \ldots, a_n) = L_n(\phi_{n-1}(a_{\pi(1)}, \ldots, a_{\pi(n-1)}), a_{\pi(n)})$$

for all $n > 2$ and some ordering rule π.

Notice that we are not imposing a unique binary operator for the whole iterative process. This was, in fact, the main criticism argued in [23] against the restrictive result obtained by Fung-Fu [13].

Right recursiveness was analogously defined (see [4]), and then the authors talk about a *recursive rule* when both left and right representations hold for the same ordering rule. We refer to *standard* recursive rules when they are based upon the identity ordering rule. Then, it follows, that a connective rule $\{\phi_n\}_{n>1}$ is recursive if and only if a set of general associativity equations hold for each n, once the ordering rule π has been already applied (see [7]):

$$\phi_n(a_1, \ldots, a_n) = R_n(a_{\pi(1)}, \phi_{n-1}(a_{\pi(2)}, \ldots, a_{\pi(n)})) = L_n(\phi_{n-1}(a_{\pi(1)}, \ldots, a_{\pi(n-1)}), a_{\pi(n)})$$

must hold for all n. A representation result was obtained in [7] by assuming certain regularity conditions.

Both associative and recursive approaches process data sequentially. This underlying assumption can be justified from an operational point of view, in case the data do not show any particular structure. Otherwise, such an underlying structure is assuming that the data is obtained according to an specific ordering (perhaps according to time, in some sense). Of course, we can analyze how the behavior of a certain pixel evolves over time, but a basic study that will focus on the comparison between each pixel and its surrounding pixels. There is no way this analysis could be properly represented according to a linear structure, at least, not within a remote sensing context, where the neighbor concept is not restricted to the pixels located in the left and right directions.

III. Classification and multicriteria procedures

From our point of view, when an image obtained by remote sensing methods is analyzed, the objective of this analysis is the obtention of meaningful classes that should capture or approximate some specific concept. We are not looking at accurate distance measures between the features of study.

In most cases, the classes are determined based on proximity measures of individual features, but a meaningful class can not be defined based on absolute measures of individual features but as a mixture of fuzzy intervals of these features. In some cases, you might find that in a particular pixel one of the features clearly indicates a high degree of membership to a class but the values of the surrounding pixels in that particular feature are all within reasonable boundaries of another class, then you have to conclude that the particular pixel, definitely belongs to the later class. This, could be the case in the example that we mentioned in the previous sections, about the house in the forest. Therefore, measurement proximity should be understood as a preprocessing stage, never as the final output. In this sense, classes may play the role of classification criteria (for example, we pursue to evaluate to what extent a certain region is a forest, which indeed is a complex concept, to be explained by a family of measured attributes).

In our search, for meaningful classes to define or explain a remote sensing image, segmentation plays a key role. Image segmentation is the part of image processing which objective is to divide the image into its constituent regions or objects. The goal is to get a manageable description of the image in terms of a partition (not necessarily crisp) into homogeneous parts that could represent the set of meaningful classes. When dealing with information from several sensors, we might have one or several three dimensional, one or several bidimensional or a mixed of bidimensional and three dimensional representations of the area being analyzed. Some segmentation procedures (see, e.g., [17], [18]) imply the aggregation of information about possible frontiers according to several directions, since in these approaches the image is defined by means of pixels arranged in a connected graph. In fact, a remote sensing image can be mathematically defined as a set P

$$P = \{p_{ij} / \quad 1 \leq i \leq r, 1 \leq j \leq s\}$$

of $r \times s$ information units -pixels-, where

$$P_{ij} = (p_{ij}^1, p_{ij}^2, \cdots p_{ij}^k)$$

is the pixel associated with the coordinate (i,j). In this way, the image is being modeled as a fuzzy planar graph (the graph is planar in the sense that two pixels p_{ij} and $p_{i'j'}$ are not connected if $\|i - i'\| + \|j - j'\| > 1$). A fuzzy graph

$$\tilde{G} = (P, \tilde{A})$$

can be then defined by the image pixels, and the set of fuzzy arcs $\tilde{A}$ will be characterized by a matrix

$$\mu_{\tilde{A}} = (\mu_{p_{ij}, p_{i'j'}})_{p_{ij}, p_{i'j'} \in \bar{P}}$$

where $\bar{P}$ is the set of connected pixels, i.e.,

$$\bar{P} = \{(p_{ij}, p_{i'j'}) \in P^2 \quad / \quad \|i - i'\| + \|j - j'\| = 1 \\ 1 \leq i \leq r, 1 \leq j \leq s\}$$

The above approach would be used as a preprocessing procedure over the data obtained from the multiple sensors. Each one of the units of the physical location of study will have a set of spatial information that has to be aggregated. The aggregation operator in this case can not be linear, the data would not only be to the right and to the left but in multiple spatial directions. The aggregation procedure should reproduce the true structure of data, so the data is successively aggregated following a path of pairs of connected units, which we anticipate should produce consistent results. We can not aggregate non connected units, since the aggregation will be meaningless.

Once the final segmentation based on an n-dimensional aggregation operator has been obtained the result can be processed by an unsupervised classification algorithm like the one designed by Del Amo *et al* ([8], [10], see also [6], [17], [18]).

Let

$$X = \{X_{i_1 j_1} / i_1 = 1, \cdots, r_1 \; j_1 = 1, \cdots, s_1\}$$

be the set of pixels to be processed (we have to remember that this information represents the output of the segmentation algorithm above), where

$$(x_{i_1 j_1}^1, x_{i_1 j_1}^2, \cdots, x_{i_1 j_1}^k) \in I\!R^n$$

is the vectorial representation of a collection of features for element $X_{i_1 j_1}$, $x_{i_1 j_1}^f$ being the value of feature f for element X_{ij}. For each one of the classes, in which the classification will be performed, a range of valid values has to be defined. The membership function for each class C_k with respect to each f property is defined following [8]. Each object $X_{i_1 j_1}$ has an associated vector

$$M_k(X_{i_1 j_1}) = (m_{1k}(x_{i_1 j_1}^1), m_{2k}(x_{i_1 j_1}^2), \cdots, m_{nk}(x_{i_1 j_1}^n))$$

for each class C_k, which shows the different degrees of verification each property has with respect to each class. The classification problem, at this point, can be solved as a multicriteria decision making problem.

In the same way as classical (crisp) sets can be defined two different but equivalent ways (by means of their extensive definition, i.e., listing those elements belonging to the crisp set, or by means of their comprehensive definition, i.e., listing those properties characterizing the crisp set), both approaches should be simultaneously be introduced in fuzzy models (see [24], [25]). Hence, when we are facing the problem of evaluating to which extent an object $x \in X$ belongs to a given fuzzy class and, we are trying to evaluate to what extent such an object verifies the properties defining such a fuzzy class. Hence, we should be answering to what extent the assertion *object x verifies those properties* is true. This has to be done by means of an existing evaluation space (which might not define a linearly ordered set of evaluation states). Within a crisp context, where one and only one evaluation state will be valid, this is being done by *choosing* one evaluation state between all elements in the evaluation space. But in a fuzzy context an object can belong to several evaluation states, with different intensities of truth or membership degrees.

Under this approach, we identify our evaluation problem as a fuzzy classification problem, so a *fuzzy classification system* (in the sense of [9]) is needed, in order to identify the degree to which each assertion *object x verifies required properties at level v* is true. As pointed out in [9], these degrees

$$\{\mu_x(c)\}_{c \in \mathcal{C}}$$

need not to sum up to one (Ruspini's partition [29] can be reached only through a specific learning procedure, in order to find an appropriate family of classes, but note that sometimes it is not possible to find a set of classes that verify Ruspini's condition, and it may not even be the objective, see [5]). Moreover, it was also pointed out in [9] that a family of classes $\mathcal{C}$ might be associated to a certain structure, not necessarily a partial order: for example, if an image is being explained by means of three basic classes *forest*, *urban area* and *water* (see [16]), it can not be properly said that *forest* is *in between urban area* and *water* (although we realize in [16] that, in some sense, the distance between *urban area* and *water* is greater than the distance between *urban area* and *forest*). Ordering is not the characteristic issue between states in a classification structure, but a relational graph, associated to a future learning of the degrees of membership.

Classes should be positively defined in order to allow evaluation learning in practice. Classifying between *tall* and *short* is in this sense correct, but classifying between *tall* and *non tall* does not seem a proper classification: *non tall* is, by definition, the negation of *tall* so we do not properly classify between these two classes: we just assign a degree of membership to *tall* and then deduce the degree of membership to *non tall* (direct estimation without direct intuition is difficult for most decision makers).

Moreover, the initial state of any learning process should be associated to an *ignorance* evaluation state, which should always be present in the evaluation space $\mathcal{C}$. In practice, $\mathcal{C}$ should be a finite and small set of evaluation states.

Then we find out that the simplest real problem will require at least two different kinds of operators: those relative to fuzzy classes in $\mathcal{C}$ (e.g., *to what extend Paul is tall and fat?*, see [24]) and those relative to crisp objects (e.g., *to what extend Paul and Mary are tall?*, see [25]). And in the same way that a particular structure should be defined in the evaluation space $\mathcal{C}$, another structure might be defined in X (see for example [17], [18], where neighborhood between pixels plays a key role).

Hence, we follow [27] claiming that, at least from a practical point of view, more attention should be devoted to type-2 fuzzy sets.

In the specific spatial classification context we are considering in this paper, we propose a classification model based upon a mapping

$$\mu : X \to [0,1]^{\mathcal{C}}$$

where

- A1 X is a well-defined non-empty, but finite, set of *objects* such that
  - A1.1 There exists a crisp directed graph (X, P), showing physical immediacy between two distinct objects $x, y \in X$ ($p_{xy} = 1$ in case there is immediacy between $x, y \in X$ and $p_{xy} = 0$ otherwise).
  - A1.2 There exists a logic on X which will allow a consistent evaluation of questions about paths of objects.
- A2 $\mathcal{C}$ is a finite evaluation space, with at least three elements, such that
  - A2.1 There is a crisp directed graph $(\mathcal{C}, R)$. ($r_{ij} = 1$ in case there is immediacy between $i, j \in \mathcal{C}$ and $r_{ij} = 0$ otherwise).
  - A2.2 There exists a logic in $\mathcal{C}$ which will allow a consistent evaluation of questions about paths of evaluation states.
- A3 There exists an *ignorance* state $I \in \mathcal{C}$ such that
  - A3.1 For every $i \in \mathcal{C}$ there exist a path connecting I with i.
  - A3.2 $\mu_x(I) = 1, \forall x \in X$ and $\mu_x(i) = 0, \forall x \in X, \ \forall i \neq I$ when there is no available information (complete ignorance).

In this way, the relative position of pixels will be taken into account in the subsequent aggregation process. Of course, see [5], additional restrictions may appear within the subsequent learning process. For example, one can expect that

$$\mu_x(I) \leq (\geq)\, \mu'_x(I)$$

should hold whenever

$$\mu_x(i) \geq (\leq)\, \mu'_x(i), \forall i$$

In the learning process the relationship between classes will play a key role.

The basic model proposed here seems to be the evaluation structure needed in many classification problems, in particular for some remotely sensed images. Meanwhile, each class can be associated to one criteria, this structure can be translated into multicriteria decision aid models, which can be upgraded by means of appropriate coloring procedures, as claimed in [14].

In particular, in [6], shows that a pure numerical search, pixel by pixel, may produce useless results, while an approach closer to the analysis of concepts produces clear improvements in the decision process. For example, the search for a concept in [6] was associated to homogeneous regions, and the objective was to obtain a set of basic concepts that would allow a good explanation of the image of study. In a first analysis, we got three general classes (*natural space*, *urban area* and *wetland*), a second analysis should be done when a more detail explanation, a higher precision is the ultimate goal for the decision maker.

Nevertheless, the aggregation of information from a set of units should be evaluated from the structure of these units. In remote sensing, for example, smoothing might be done in a first analysis just computing, for each pixel, the aggregated value of the immediate contiguous pixels, or could included influence values of a larger number of pixels around (expanded neighborhood). This approach implies an aggregation in several phases or stages. At each stage we will add one extra outer line around the pixel, with all those pixels being equidistant with respect to the selected pixel. Once this process has been performed we will aggregate the aggregated value of every outer line (depending on our objective, the difference between each step of the incremental analysis can be considered, and also the distance between the pixel of study and the aggregation of all those pixels in the outer line around it). The aggregation rule to be used within an outer line is expected to be commutative, since pixels at the same level are considered equally important. In case we are considering multiple layers of pixels the aggregation rule will be weighted, assigning more weight to the lines closer to the pixel under study. An analogous situation appears when the objective is detecting borders instead of smoothing the area of study. As proposed in [17], for example, we can search for borders in different directions, developing an aggregation procedure for each direction, individually. The decision about a border pixel would be made only based on the computation of an aggregated value of all those directions. As a conclusion, the aggregation rule through the incremental layers is expected to be different that the aggregation of the aggregated values in each one of the directions of study. Different data structures will lead to different aggregation procedures.

IV. Final remarks

There is a general interest to design innovative techniques to improve the decision making process via fusion of information, particularly, in homeland security applications, including situational awareness, precision strike, etc. The data acquisition capabilities based on the sensor utilization has grown to extremes were the amount of data collected needs to be automatically analyzed. The amount of data collected for the same phenomena fits the description of the typical set of data that should be automatically analyzed and aggregated by means of multicriteria decision making techniques.

In this paper we proposed to bring a particular model, initially conceived for classification, into a multicriteria decision making context. This translation can be useful when the classes we are trying to determine play in fact the role of criteria, which in general imply fuzzy concepts. Our purpose then is to take advantage of previous research done by the authors concerning segmentation and other representation techniques. These representation techniques are based upon coloring of fuzzy classes, which can be understood as a first stage for multicriteria models based on fuzzy set theory.

Acknowledgment

This research has been partially supported by Smiths Aerospace as an $IR\&D$ project and by the Government of Spain, research grant TIN2006-06190.

References

[1] R.E. Barlow and F. Proschan, *Statistical theory of reliability and life testing*. To Begin With, Silver Spring, 1981.

[2] K.Y. Cai, *Introduction to Fuzzy Reliability*. Kluwer, Boston, 1996.

[3] T. Calvo, G. Mayor and R. Mesiar, *Aggregation Operators.* Physica-Verlag, Heidelberg 2002.

[4] V. Cutello and J. Montero: "Recursive connective rules". *Intelligent Systems* 14:3–20, 1999.

[5] A. Del Amo, D. Gómez, J. Montero and G. Biging, "Relevance and redundancy in fuzzy classification systems," *Mathware and Soft Computing* 8:203–216, 2001.

[6] A. Del Amo, D. Gómez, J. Montero and G. Biging, "Improving fuzzy classification by means of a segmentation algorithm." In H. Bustince, F. Herrera and J. Montero, eds., *Fuzzy Sets and Their Extensions*, Springer, in press.

[7] A. Del Amo, J. Montero and E. Molina, "On the representation of recursive rules, " *European Journal of Operational Research* 130, 29–53, 2000.

[8] A. Del Amo, J. Montero and G. Biging, "Classifying pixels by means of fuzzy relations," *International Journal of General Systems* 29:605–621, 1999.

[9] A. Del Amo, J. Montero, G. Biging and V. Cutello, "Fuzzy classification systems," *European Journal of Operational Research* 156, 459-507, 2004.

[10] A. Del Amo, J. Montero, A. Fernandez, M. Lopez, J. Tordesillas and G. Biging, "Spectral fuzzy classification: an application," *IEEE Trans on Systems, Man and Cybernetics, Part C* 32:42–48, 2002.

[11] J. Dombi, "Basic concepts for a theory of evaluation: the aggregative operator," *European Journal of Operational Research* 10:282–293, 1982.

[12] G.M. Foody, "The continuum of classification fuzziness in thematics mapping," *Photogrammetric Engineering and Remote Sensing* 65:443–451, 1999.

[13] L.W. Fung and K.S. Fu., "An axiomatic approach to rational decision making in a fuzzy environment"; in L.A. Zadeh *et al.*, eds.: *Fuzzy sets and their applications to Cognitive and decision processes*, Academic Press, New York; pp. 227–256, 1975.

[14] D. Gómez, V. López and J. Montero, "The role of fuzziness in decision making". In D. Ruan, E. Kerre and P. Wang, eds., *Fuzzy Logic: an spectrum of applied and theoretical issues*, Springer, in press.

[15] D. Gómez and J. Montero, "A discussion on aggregation operators," *Kybernetika* 40, 107–120, 2004.

[16] D. Gómez, J. Montero and G. Biging: Accuracy measures for fuzzy clasification. *Proceedings Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU Conference*, Paris, France; pp. 1556–1563, 2006.

[17] D. Gómez, J. Montero, J. Yáñez and C. Poidomani, "A graph coloring algorithm approach for image segmentation," *OMEGA-Int.J.Manag.Sci.* 35, 173–183, 2007.

[18] D. Gómez, J. Montero and J. Yanez, "A coloring fuzzy graph approach for image classification," *Information Sciences* 176:3645-3657, 2006.

[19] J. R. Jensen, *Introductory Digital Image Processing. A Remote Sensing Perspective*. Prentice Hall, 1986.

[20] E.P. Klement, R. Mesiar and E. Pap, *Triangular norms*. Kluwer, Dordrecht, 2000.

[21] G.J. Klir and T.A. Folger, *Fuzzy sets, uncertainty and information*. Prentice Hall, London, 1988.

[22] K. Menger, *Statistical methods*. Proc. Nat. Acad. Sci. U.S.A. 8:535–537, 1942.

[23] J. Montero, "A note on Fung-Fu's theorem," *Fuzzy Sets and Systems* 17:259–269, 1985.

[24] J. Montero, "Comprehensive fuzziness," *Fuzzy Sets and Systems* 20:89–86, 1986.

[25] J. Montero, "Extensive fuzziness," *Fuzzy Sets and Systems* 21:201–209, 1987.

[26] J. Montero, "Aggregation of fuzzy opinions in a fuzzy group," *Fuzzy Sets and Systems* 25, 15–20, 1988.

[27] J. Montero, D. Gómez and H. Bustince, "Atanassov's intuitionistic fuzzy sets as a classification model." *Proceedings International Fuzzy Systems Association Conference*, Cancún, México, 18-21 June, 2007.

[28] J. Montero, J. Tejada and J. Yáñez, "Structural properties of continuous systems," *European Journal of Operational Research* 45, 231–240, 1990.

[29] E.H. Ruspini, "A new approach to clustering," *Information and Control* 15:22–32, 1969.

[30] B. Schweizer and A. Sklar, *Probabilistic metric spaces*. North Holland, Amsterdam, 1983.

[31] R.R. Yager, "On ordered averaging aggregation operators in multicriteria decision making," *IEEE Transactions on Systems, Man and Cybernetics* 18:183–190, 1988.

[32] R.R. Yager and A. Rybalov, "Uninorm aggregation operators," *Fuzzy Sets and Systems* 80:111–120, 1996.