Towards Interpretable General Type-2 Fuzzy Classifiers

Luís A. Lucas Tania M. Centeno
Myriam R. Delgado
Federal University of Technology – Paraná.
Av. Sete de Setembro, 3165. Curitiba - PR, 80230-901, Brazil
{lalucas, mezzadri, myriamdelg}@utfpr.edu.br *

Abstract

This paper presents two versions of a general type-2 fuzzy classifier. The focus is on interpretability since the rules are meaningful and the rule base is comprised of few rules, which is a direct consequence of the hierarchical reclassification process being proposed. The approaches are evaluated on a land cover classification problem by using data from a remote sensing platform. The classifiers' performance are compared with the reference ones' (maximum likelihood classifier and ordinary fuzzy classifier). The results show that the general type-2 fuzzy modeling is able to produce accurate classifiers while maintaining the model interpretability.

1. Introduction

Fuzzy rule-based models may be considered a linguistic representation of the mental model of a certain system by means of experience [1]. When the knowledge required is not easily available to design fuzzy systems the development of computer techniques to extract and represent knowledge in a fuzzy rule-based system may be necessary. Data-driven fuzzy modeling, or fuzzy modeling (FM) for short, has attracted interest of many researches where model interpretability plays a central role [1, 4]. When the focus is on interpretability, linguistic fuzzy modeling must generate systems for which the language is easily interpretable by human beings and compact fuzzy rule bases with short individual rules are essential conditions.

The multisource classification model relies on input-by-input classification techniques that combine different sources of information (e.g., multisensor data, multiband data, symbolic data such as model-based knowledge represented by if-then rules, etc.).

*Myriam Delgado acknowledges the CNPq grant 307735/2008-7 and Fundação Araucária process N. 233/07 - 8331; Tania Centeno the CNPq grant 304867/2008-0, for their support.

Recently, several applications on type-2 fuzzy inference systems appeared. Such systems are interesting because they can handle both sorts of uncertainty: the one presented when measuring a signal (i.e., noise) and the other at the semantic level (linguistic terms can vary from one expert to another) [20]. Type-2 fuzzy inference systems have great potential to be employed in digital classification [14] (specially in multisource classification) but most applications use interval type-2 sets [12, 13, 17, 18] which are however simplifications of general type-2 ones having the lower computational cost as an advantage.

In this paper, two versions of a general type-2 fuzzy classifiers are presented: specific and general rule-based systems, where the last one enables don't care conditions in the rule antecedent. Both classifiers are based on previous work where we established a type-2 inference mechanism referred to as "General Type-2 Fuzzy Scaled Inference" [8] and on type-2 fuzzy classifiers [9, 10]. The aim here is on interpretability aspects. Because of the hierarchical reclassification process being proposed, this data-driven fuzzy type-2 modeling is able to keep small rule bases (the total of fuzzy rules is equal to the total of classes), each rule being comprehensible, while maintaining the system accuracy. The classifiers' performance will be compared against the maximum likelihood classifier and the ordinary fuzzy classifier (type-1 fuzzy classifier).

2 Type-2 Fuzzy Sets

Type-2 fuzzy sets were first presented by Zadeh as an evolution for his theory of ordinary fuzzy sets [19] but, later on, they were studied deeply and their theory basis was appropriately established by Mendel [6]. Nowadays, the most accepted definition stays that type-2 fuzzy sets are those ones whose membership grades are themselves ordinary (type-1) fuzzy sets. Type-2 sets which strictly follow the given definition are called general type-2 fuzzy sets and those that employ constant membership functions (unitary ones) are known as interval type-2 fuzzy sets.

978-0-7695-3872-3/09 $26.00 © 2009 IEEE
DOI 10.1109/ISDA.2009.28

Thus a general type-2 fuzzy set $\tilde{A}$ can be defined at the universe X as

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} \mu_{\tilde{A}}(x, u)/(x, u), \quad (1)$$

where $J_x \subseteq [0, 1]$ is the set of primary membership grades of $x \in X$, with $u \in J_x$, $\forall x \in X$, and $\mu_{\tilde{A}}(x, u)$ is the type-2 membership function [16].

According to [5], for a particular $x' \in X$, Equation (1) becomes

$$\mu_{\tilde{A}}(x = x', u) \equiv \mu_{\tilde{A}}(x') = \int_{u \in J_{x'}} f_{x'}(u)/u \quad (2)$$

where $\mu_{\tilde{A}}(x')$ is called the vertical slice of $\mu_{\tilde{A}}(x, u)$ when $x = x'$ [5], $J_{x'} \subseteq [0, 1]$ is the set of primary membership grades of x' and $f_{x'}(u)$, $0 \leq f_{x'}(u) \leq 1$, is a function f of the primary membership grade u, that identifies the secondary membership grades of x' in $\tilde{A}$.

In General Type-2 Fuzzy Sets (GT2 FS), the primary memberships whose secondary grades are equal to one are called principal memberships. There is a particular, but important, case that, for each x_i, there exists only one principal membership, which is called Principal Membership Function (PMF) [6, 14]. PMF is thus given by Equation (3).

$$\mu_{\tilde{A}_{princ}} = \mu_A = \begin{cases} u & \forall x \mid f_x(u) = 1 \\ 0 & otherwise \end{cases} \quad (3)$$

For an example of a general type-2 fuzzy set consider Figure 1 which was drawn with the aid of the *General Footprint of Uncertainty* (GEFOU), where the intensity of shading is proportional to secondary membership grades [8]. Figure 1(a) shows the GEFOU of a general type-2 fuzzy set $\tilde{A}$, whose Gaussian PMF is represented by the thin white line and Figure 1(b) shows a vertical slice of $\tilde{A}$ at $x = x'$.

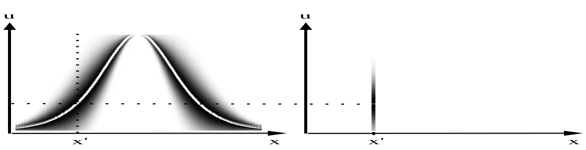


Figure 1. a) A General type-2 Fuzzy Set; b) Its corresponding vertical slice at $x = x'$.

3 The Proposed Classifier

In this paper we consider a General Type-2 Fuzzy Classifier (GT2 FC) whose typical structure [7, 10] is illustrated in Figure 2.

In this work, the GT2 FC has p input variables $x_1 \in X_1$, $x_2 \in X_2, \cdots, x_p \in X_p$, and one output $y \in Y$.

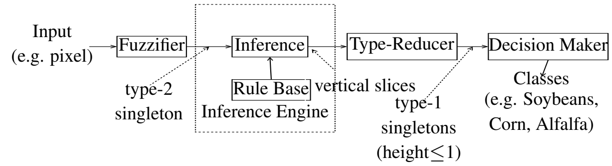


Figure 2. Type-2 fuzzy classifier's inference mechanism.

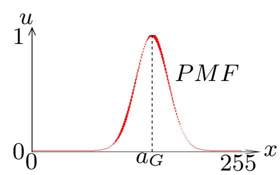
Suppose that the classifier's rule base is composed of L rules, whose l_{th} rule is given by Equation (4).

$$R^l : \text{IF } x_1 \text{ is } \tilde{F}_1^l \; op \; \cdots \; op \; x_p \text{ is } \tilde{F}_p^l, \text{ THEN } y \text{ is } \tilde{G}^l; \quad (4)$$

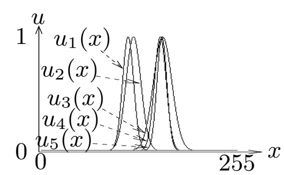
where the T2 FS $\tilde{F}_1^l$, $\tilde{F}_2^l \cdots \tilde{F}_p^l$ are the rule's antecedents (type-2 fuzzy sets) and $\tilde{G}^l$ is the rule's consequent (a type-2 singleton [10]). The term op means the antecedent's aggregation operator (e.g., AND, OR).

Each antecedent $\tilde{F}_i^l$ can be built based on its Principal Membership Function (PMF). Admit a $\tilde{F}_i^l$ with a Gaussian PMF given by $PMF = e^{\frac{-(x-a_G)^2}{2 \cdot \sigma_G^2}}$, where a_G is the global average of input values x considering all training samples of a given class taken together, as a big sample, and σ_G is the global standard deviation of these data around a_G.

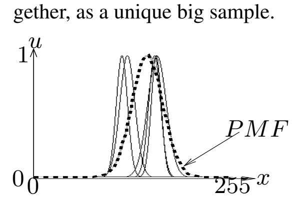
Figure 3 shows the construction of a type-2 antecedent from five training samples (five experts). In Figure 3(a), it is shown the resulting PMF (all the inputs of all the samples for a given class are taken together).



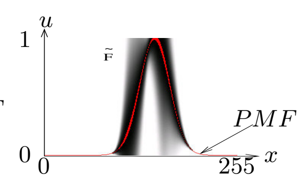
(a) PMF obtained from all the samples of a class taken together, as a unique big sample.



(b) Gaussian functions which describe each sample (expert).



(c) Comparison between PMF, dashed line, and each sample, full line.



(d) Final type-2 fuzzy set.

Figure 3. Construction of a type-2 fuzzy set from training samples.

To obtain the type-2's secondary grades we suggest to take into account possible differences between the samples of the same class (where each sample may represent a specific expert). Thus, trying to capture such differences, we propose that the training samples are taken individually and from each of them we can build a type-1 fuzzy set (a Gaussian type for example). So, the Gaussian $u_j(x) = e^{\frac{-(x-a_j)^2}{2\cdot\sigma_j^2}}$, $j = 1 \cdots J$, may represent the j^{th} sample of a given class , where J is the number of training samples, a_j is the average and σ_j is the standard deviation of sample j. Figure 3 (b) shows the Gaussian functions, one for each sample.

Thereafter we build the vertical slices $\mu_{\tilde{F}}(x_k)$ at each coordinate x_k, taking into account the dispersion $\sigma(x_k)$, which exists between the samples $u_j(x)$ at x_k, i.e., $\mu_{\tilde{F}}(x_k) = e^{\frac{-(u(x_k)-PMF(x_k))^2}{2\cdot\sigma^2(x_k)}}$, where $PMF(x_k)$ is the value of the PMF in $x = x_k$ and $u(x_k)$ being the primary membership grades of the vertical slice $\mu_{\tilde{F}}(x_k)$ also at $x = x_k$. The total dispersion of the training data, $\sigma(x_k)$, can be calculated by the standard deviation of the Gaussian $u_j(x_k)$ at coordinate x_k.

In Figure 3 (c), we can see a comparison between the PMF, in dashed line, against the Gaussians, in full lines, and, in (d), the resulting type-2 set. We can see that, when we approximate to the universe limits ($x = 0$ and $x = 255$) secondary grades tend to be defined solely by the PMF. This is because the dispersion among Gaussian functions $u_j(x)$ is null in these points.

3.1 Matching

At the inference process performed by the proposed classifier, the first step comprises the *matching* between the i^{th} fact ($\tilde{F}_i$ representing the fuzzification applied over the i^{th} input) and the i^{th} antecedent that appears at l^{th} rule ($\tilde{F}_i^l$ representing the knowledge of the i^{th} source of information at l^{th} rule). A *matching* is a sort of similarity measure between two fuzzy sets (type-2 sets in our case) and it is calculated between an input and its corresponding antecedent.

According to [8] the *matching* $\tilde{M}_i^l$ between the input $\tilde{F}_i$ (the fact) and the antecedent $\tilde{F}_i^l$ (the premise) can be calculated by the generalized *sup-t* operation [7], $\tilde{M}_i^l = sup\left[\tilde{F}_i \cap \tilde{F}_i^l\right]$, which by itself is comprised of two operations: the intersection between an input and its corresponding antecedent ($\tilde{I}_i^l = \tilde{F}_i \cap \tilde{F}_i^l$) and the supremum of the resulting type-2 fuzzy set $\left(sup\left[\tilde{I}_i^l\right]\right)$. Both operations are explained in the following.

As described in [8], the intersection $\tilde{I}_i^l = \tilde{F}_i \cap \tilde{F}_i^l$ between the input $\tilde{F}_i$ and its corresponding antecedent $\tilde{F}_i^l$ can be obtained by Equation (5) as

$$\tilde{F}_i \cap \tilde{F}_i^l = \int_{x\in X}\int_u\int_w f_x(u) \wedge g_x(w)/u \wedge w/x \qquad (5)$$

where u, $w \in J_x \subseteq [0,1]$ are the primary membership grades of $\tilde{F}_i$ and $\tilde{F}_i^l$, and $f_x(u)$ and $g_x(w)$ are the secondary membership grades of $\tilde{F}_i$ and $\tilde{F}_i^l$, respectively, $\wedge$ is a t-norm and $\vee$ is a t-conorm. In this paper we adopted *min* as t-norm and *max* as t-conorm.

According to [8] the supremum $\tilde{M}_i^l$ of type-2 fuzzy set $\tilde{I}_i^l$, that is defined at $x = x'$, can be calculated by

$$\tilde{M}_i^l = sup\left[\tilde{I}_i^l\right] = \left[\bigsqcup_{x\in X}\mu_{\tilde{I}_i^l}(x)\right]/x' =$$
$$= trs_{\ x=x'}\left[\mu_{\tilde{I}_i^l}(x_1)\right] \sqcup \cdots \sqcup trs_{\ x=x'}\left[\mu_{\tilde{I}_i^l}(x_N)\right] \qquad (6)$$

where $\mu_{\tilde{I}_i^l}(x_1)\cdots\mu_{\tilde{I}_i^l}(x_N)$ are the vertical slices of type-2 fuzzy set $\tilde{I}_i^l$, $trs_{\ x=x'}\left[\mu_{\tilde{I}_i^l}(x_j)\right]$ is the translation of the vertical slice $\mu_{\tilde{I}_i^l}(x_j)$, defined at $x = x_j$, into the new position $x = x'$ and $\sqcup$ is the *join* operator [15].

Figure 4 shows an example of *matching* $\tilde{M}_i^l$ (enlarged in Figure 4c to improve visualization) from the intersection $\tilde{I}$ (depicted in Figure 4b) of a Gaussian fuzzy input $\tilde{F}_i$ and its corresponding antecedent $\tilde{F}_i^l$ (both depicted in Figure 4a).

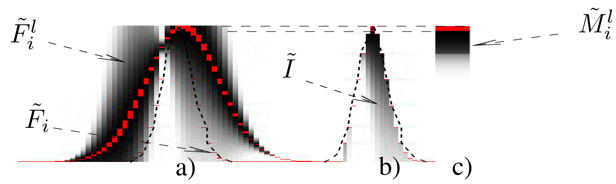


Figure 4. An example of matching $\tilde{M}_i^l$ (in c) obtained from the intersection $\tilde{I}$ (in b) calculated from input $\tilde{F}_i$ and antecedent $\tilde{F}_i^l$ (in a).

3.2 Antecedent's aggregation: *AND*,*OR*

After the *matching*, the next step is the antecedents aggregation. It is possible to aggregate antecedents by the *AND* and *OR* operators in Equation (4). When aggregating antecedents by *AND*, the process is very restrictive, and in this case, if only one source of information disagrees (resulting in a null matching) the rule becomes inactive.

Being less restrictive, it is possible to aggregate antecedents by the OR operator, which does not produce a zero rule's output when a null matching appears in it. This is because it requires only one not-null matching to produce a non-zero rule (the OR operator implemented by *max* considers the highest matching when inferring the rule's output). The disadvantage of OR operator is that it takes solely one source of information into account (the one with the highest matching) so OR rule-based systems are often less accurate than the ones produced by *AND* operator (where all sources of information are taken into account).

To deal with restrictive *versus* permissive characteristics of *AND* and *OR* operators, two different rule bases are being proposed and will be detailed in the next section.

3.3 Specific *versus* General Rule Bases

Specific fuzzy rules assume all their antecedents to be useful. Thus they take all their corresponding inputs into account when performing the inference process. In data-driven fuzzy modeling it is important to deal with feature selection, so some fuzzy inference systems employ general fuzzy rules [2].

In such systems, it is possible to include "don't care" conditions in the rule's antecedents i.e., not all inputs are relevant to the rule's output inferring process.

In this paper, we are presenting two versions of general type-2 fuzzy classifier: "Specific Rule Base fuzzy classifier with *AND*, followed by OR, antecedent's aggregation operator (SRB-AND/OR)" and "General Rule Base fuzzy classifier with *AND* rule aggregation operator (GRB-AND)".

The first type of classifier (the one that uses the specific rule base) always uses all the available sources of information (antecedents), firstly by using the *AND* aggregation operator and, if all the rules are inactive, the input is reclassified considering the OR operator in the antecedent aggregation. The assumption here is that, in absence of input outliers, the classification taking P sources of information (SI) into account is more confident than the one taking $P-k$ SI, $k = 1 \cdots P-1$: the more SI taking part in the decision, the more accurate the classification is expected to be. The proposed classifier performs at most two cycles of classification for each input pattern:: the first one with *AND* and the second one, if necessary (because the first one was not successful), with *OR*.

The second version of the classifier (the one that uses the general rule base) assumes that not all the conditions must be considered to infer the output (a kind of feature selection). Besides resulting in more compact rule bases the exclusion of some linguistic terms can produce shorter fuzzy rules, improving the model interpretability. GRB-AND performs at most P cycles of classifications for each input, where P is the number of sources of information. All these classifications are performed by means of *AND* operator. The first one is tried with all antecedents but if no rule is activated because of null matching, one antecedent from each rule (the one that caused the null matching) is taken off and the input is reclassified with $P-1$ sources of information. The process proceeds till the input can be classified, even by only one source of information. If classification cannot be done, after P trials, the label NC (non classified) is assigned to the input.

The pseudocode summarizes the steps performed by GRB-AND system

Hierarchical Reclassification (SRB-AND)

```
for input i to total_size
  n.antecedents ← total of SI
  evaluate input i by type-2 inference
  while (no activation and n.antecedents>1)
      n.antecedents ← n.antecedents-1
      evaluate input i by type-2 inference
  end while
  if (no activation or conflicting classes)
       error
  else
       input classified (correctly or not)
  end if
end for
```

Pseudocode 1

Although both approaches (SBR-AND/OR and GRB-AND) have been proposed to deal with null matching, they can still appear resulting in non-classified inputs, as will be discussed in the next section.

3.4 Measure of Model accuracy

There are two situations where classification is not possible (NC — non-classified input): 1) excess of null matching ($\tilde{I} = \tilde{F}_i \cap \tilde{F}_i^l = \phi$, $l = 1, \cdots, L$); 2) confusion, where more than one rule are activated with different class label and the same output level i.e., the input resembles 2 or more classes ("mixed input"). Null matchings are quite normal in a fuzzy system but when they occur all over the rule base (no rule is activated) after all the cycles of classification, they can be considered an error of classification caused by gaps in the input universes. Confusion may also occur, but in this paper it will be considered as an error.

The measure of model accuracy considers these two types of errors (caused by non-classified pixels) and is given by the expression $Prec = \sum_{k=1}^{N} CC(k)/N$, where $Prec$ is the precision level, $CC(k)$ is *one* if the k^{th} input pattern is correctly classified and *zero* otherwise, and N is the total of testing patterns considered at each cycle of classification.

4 Case Study

For the case study we consider the problem of land cover classification. In this particular subject, the sources of information are spectral bands. Classes become thematic ones and inputs are pixels (singleton fuzzification) or group of pixels (fuzzy set fuzzification).

The problem of "mixed input" becomes a "mixed pixel" one: depending on the remote sensing platform we are getting information from (e.g., Landsat), one pixel could be related to big areas at ground level (e.g., 30×30m) so it is quite common that a pixel corresponds to mixed agricultural crops (more than one thematic class in one pixel). Moreover, crops at different geographical locations can be at different growth stages meaning that they can appear somewhat different when observed from the space.

The data set is comprised of twelve sources of information (spectral bands) with resolution of 220×949 pixels and was gathered by an airborne scanner which flew the southern part of Tippecanoe County, Indiana, United States, in June 1966 [11]. Each spectral sensor responds in a different wavelength such that we possess data ranging from 400nm to 1000nm organised in the following way: band 1 (B1) works from 400nm to 440nm, B2 440 - 460 nm, B3 460 - 480 nm, B4 480 - 500 nm, B5 500 - 520 nm, B6 520 - 550 nm, B7 550 - 580 nm, B8 580 - 620 nm, B9 620 - 660 nm, B10 660 - 720 nm, B11 720 - 800 nm and B12 800 - 1000 nm.

Figure 5 presents the ground truth for the considered region where we can see that white areas are those cultivated with corn, in light gray we have soybeans, in dark gray wheat and in black, we have the image background [11].

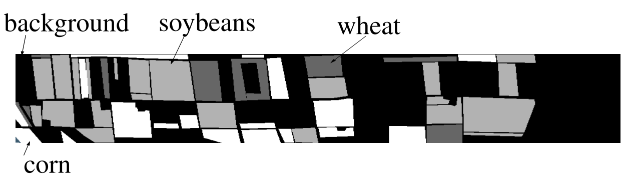


Figure 5. Ground truth for the data set.

We will employ all the sources of information to test the two proposed classifiers (GRB-AND and SRB-AND/OR) and the resulting accuracies will be compared against type-1 and ML classifiers.

4.1 Evaluation criterion

We performed a cross validation similar to that suggested by [3], employing all the available sources of information (i.e., 12 spectral bands). In this way, seven squared evaluating samples with 81 pixels (9×9) were randomly chosen resulting in 567 (7×81) pixels. At each turn 6 samples (486 pixels) are used to train the classifier (i. e. to construct the antecedents $\tilde{F}_i^l$) and one sample (81 pixels) is used to test the classifier.

In the proposed method, when building the rule base, each antecedent $\tilde{F}_i^l$ is obtained from band i, $i = 1 \cdots P$, in such a way that the number of antecedents (variables) is equal to the number of considered sources of information (12 in our case).

In the land cover classification context addressed here, the design phase considers the sources of information and selects 6 samples (out of the 7 randomly chosen) as training ones to build the type-2 fuzzy set (T2 FS) that best represents each class at each spectral band.

Figure 6 shows the evaluating samples we will employ: those labeled $C_0 \cdots C_6$ are for corn, $S_0 \cdots S_6$ for soybeans and $W_0 \cdots W_6$ for wheat.

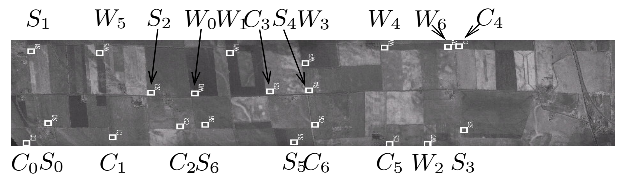


Figure 6. Evaluation samples.

Because we work with 3 classes and 7 samples per class, it would be necessary 7×7×7=343 times so that all evaluation samples, from all classes, have the opportunity to become testing samples. The accuracy of all turns are averaged. The same will be done for each of the 4 classifiers: GRB-AND, SRB-AND/OR, ML and T1.

5 Results and Discussion

The experiment suggested at section 4.1 was performed and the results (measured by means of a seven-fold cross-validation) are presented at Table 1. It is possible to rank the classifiers from the best to the worst: GRB-AND, SRB-AND/OR, T1 and ML.

Table 1. Classifiers' accuracies after the cross-validation.

	ML	**T1**	**SRB-AND/OR**	**GRB-AND**
Avg. acc.	0.8120	0.8398	0.8715	0.8747
Std Dev	0.1287	0.1247	0.1257	0.1264
Min	0.4400	0.5000	0.5226	0.5200
Max	0.9100	1.0000	1.0000	1.0000

It can be seen that, on average, the best classifier (the type-2 GRB-AND) correctly classified 87,47% of the applied inputs and that the worst classifier (ML) correctly classified 81,20% of the inputs.

Kruskal/Wallis and Anova tests were performed over the classifiers' accuracies obtained at the 343 testing turns (section 4.1) to evaluate if differences presented in Table 1 are statistically significant. We employed all the sources of information. We could see that the type-2 fuzzy classifiers are statistically better than the ordinary fuzzy classifier and the ML one, but there is not statistical difference between specific (SRB-AND/OR) and general rule bases (GRB-AND).

In spite of building data-driven fuzzy classifiers, we were successful on keeping the number of rules small (only 3 rules). Large rule bases contribute to producing incomprehensible fuzzy reasoning so we tried, as much as possible, to keep the number of rules small. The total number of rules (L) that encompass the rule base is equal to the number of classes i.e., there is one rule for each class, and this is independent of the number of sources of information. This was possible due to the hierarchical classification that tries to resolve universe's uncovered regions by reclassifying such regions (inputs) as described in section 3.3. It should be pointed out that in some particular situations (e.g. XOR classification problem), it will be necessary more than one rule per class.

It is not enough to have a small rule base if the rules themselves are messy. So we work out rules that are both simple and comprehensible. To achieve these goals we designed the classifier in such a way that there were as many antecedents as the number of available sources of information and that there is only one consequent per rule i.e., a type-2 singleton [10] representing the desired label. Also, rules can be read in an intelligible straightforward manner. An example of individual rule produced by the classifier GRB-AND is: "If source of information 1 (SI1) says class 1 and source of information 4 (SI4) says class 1 then pixel will be classified as class1". As can be noticed, this short (the remaining SI have been considered don't care) and simple rule could be easily interpretable by human beings.

6 Conclusions and future work

We could show that, for the employed data set, both versions of type-2 classifiers (GRB-AND and SRB-AND/OR) were better than the type-1 counterpart and the ML statistical classifier.

In this paper we could also show that non-classified inputs could be reclassified hierarchically and that it was possible to build classifiers with few rules, producing a comprehensible system where the number of rules is small and the rules themselves are simple and straightforward.

In the future we intend to compare our general type-2 classifiers against an interval type-2 one. We should investigate the hypothesis that the simplification of the type-2 set (general → interval) gives rise to accuracy decrease.

References

[1] J. Espinosa and J. Vandewalle. Constructing fuzzy models with linguistic integrity from numerical data-AFRELI algorithm. *IEEE Transactions on Fuzzy Systems*, 8(5):591–600, 2000.

[2] H. Ishibuchi, T. Nakashima, and T. Murata. Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems. *IEEE Transactions on Systems, Man, and Cybernetics — Part B: Cybernetics*, 29(5):601–618, 1999.

[3] H. Ishibuchi and T. Yamamoto. Rule weight specification in fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems*, 13(4):428–435, 2005.

[4] Y. Jin. Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement. *IEEE Transactions on Fuzzy Systems*, 8(2):212–221, 2000.

[5] R. John, J. Mendel, and J. Carter. The extended sup-star composition for type-2 fuzzy sets made simple. In *IEEE International Conference on Fuzzy Systems*, pages 7212–7216, Vancouver, Canada, July 2006.

[6] N. K. Karnik and J. M. Mendel. Operations on type-2 fuzzy sets. *Fuzzy Sets and Systems*, 122:327–348, 2001.

[7] N. K. Karnik, J. M. Mendel, and Q. Liang. Type-2 fuzzy logic systems. *International Journal on Fuzzy Systems*, 7(6):643–658, December 1999.

[8] L.A.Lucas, T. M. Centeno, and M. R. Delgado. General type-2 fuzzy inference systems: Analysis, design and computational aspects. In *IEEE International Conference on Fuzzy Systems*, pages 1107–1112, London, UK, jul 2007.

[9] L.A.Lucas, T. M. Centeno, and M. R. Delgado. General type-2 fuzzy classifiers to land cover classification. In *ACM Symposium on Applied Computing*, pages 1743–1747, Ceará, Brazil, mar 2008.

[10] L.A.Lucas, T. M. Centeno, and M. R. Delgado. Land cover classification based on general type-2 fuzzy classifiers. *International Journal of Fuzzy Systems*, 10(3):207–216, set 2008.

[11] D. A. Landgrebe. Multispectral data analysis: A moderate dimension example. West Lafayette, Jan 1997.

[12] Q. Liang and J. M. Mendel. Interval type-2 fuzzy logic systems. In *IEEE International Conference on Fuzzy Systems*, pages 328–333, May 2000.

[13] Q. Liang and J. M. Mendel. Interval type-2 fuzzy logic systems: theory and design. *International Journal on Fuzzy Systems*, 8(5):535–550, 2000.

[14] Q. Liang and J. M. Mendel. MPEG VBR video traffic modeling and classification using fuzzy technique. *IEEE Transactions on Fuzzy Systems*, 9(1):183–193, February 2001.

[15] J. M. Mendel. *Uncertain Rule-Based Fuzzy Logic Systems*. Prentice Hall PTR, 2000.

[16] J. M. Mendel and R. I. B. John. Type-2 fuzzy sets made simple. *IEEE Transactions on Fuzzy Systems*, 10(2):117–127, April 2002.

[17] J. M. Mendel and F. Liu. Super-exponential convergence of the karnik-mendel algorithms used for type-reduction in interval type-2 fuzzy logic systems. In *IEEE International Conference on Fuzzy Systems*, pages 6603–6610, Vancouver, Canada, July 2006.

[18] H. Wu and J. Mendel. Classification of battlefield ground vehicles using acoustic features and fuzzy logic rule-based classifiers. *IEEE Transactions on Fuzzy Systems*, 15(1):56–72, 2007.

[19] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning-I. *Information Sciences*, 8:199–249, 1975.

[20] H. Zeng and Z.-Q. Liu. Type-2 fuzzy sets for pattern recognition: The state-of-the-art. *Journal of Uncertain Systems*, 1(3):163–177, 2007.