



HAL
open science

Fuzzy rule classifier: Capability for generalization in wood color recognition

Vincent Bombardier, Emmanuel Schmitt

► **To cite this version:**

Vincent Bombardier, Emmanuel Schmitt. Fuzzy rule classifier: Capability for generalization in wood color recognition. *Engineering Applications of Artificial Intelligence*, 2010, 43 (6), pp.978-988. 10.1016/j.engappai.2010.05.001 . hal-00504735

HAL Id: hal-00504735

<https://hal.science/hal-00504735v1>

Submitted on 16 May 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fuzzy Rule Classifier: Capability for generalization in Wood Color Recognition

Vincent Bombardier^a, Emmanuel Schmitt^a

^aResearch Centre for Automatic Control – CRAN, CNRS UMR 7039
Université Henri Poincaré – Campus scientifique, BP 239
54506 Vandoeuvre-lès-Nancy Cedex – France
vincent.bombardier@cran.uhp-nancy.fr, emmanuel.schmitt@cran.uhp-nancy.fr

Abstract

In this paper, a classification method based on fuzzy linguistic rules is exposed. It is applied for the recognition of the gradual color of wood in an industrial context. The wood, which is a natural material, implies uncertainty in the definition of its color. Moreover, the timber context leads obtaining imprecise data. Several factors can have an impact on the sensors (ageing of the acquisition system, variation of the ambient temperature, etc.). Finally, the data sets are often small and incomplete. Thus the proposed method must work within these constraints, and must be compatible with the time-constraint of the system. This generally imposes a weak complexity of the recognition system. The Fuzzy Rule Classifier is split in two main parts, the fuzzification step and the rule generation step. To improve the tuning of this classifier, a specific fuzzification method is presented and compared with more classical ones. Several comparisons have been made with other classification method such as neural network or support vector machine. This experimental study showed the suitability of the proposed approach essentially in term of generalization capabilities from small data sets, and recognition rate improvement.

Keywords: Classification, Fuzzy Logic, Image Processing, Fuzzy Rules, Color Recognition.

1. Introduction

The works presented in this article deals with color classification on wooden boards in an industrial environment. Color recognition is an important step for matching wood pieces which affects many activities of the timber industry (veneer, paneling, manufacture of squares, etc.). The wood pieces have to be assembled according to their perceived color. The aim is mainly to provide a wood piece that seems to be homogeneous and massive. This kind of problems belongs to the aesthetic aspect of wood characterization. The aspect control issue can be divided into two categories:

- detection, localization and identification of singularities and
- classification and identification of color and / or grain (textured aspect of wood).

In the context of wood classification (color, texture and singularities), different methods can be found in the literature. For the defect recognition, there are essentially “compilation” methods which “compile” an important training data set to obtain the output classes’ model: Neural Networks (NN) [1-3] and Genetic Algorithms (GA) [4]. For color or texture identification, the used methods are principally based on Neural Networks (NN) [5-7] but also on k Nearest Neighbor algorithm (k-NN) [8] [9]. Distance Minimization algorithms [10 – 13] or Genetic Algorithms [14] are used too.

The aim of the global vision system is to identify the wood color in a continuous mode during the production. Such a system involves lots of constraints.

Firstly, the classification method must be able to work with small training data sets. Indeed, some classes which are rare in nature are defined with few samples. Moreover, providing a training data set requires the industrialist to label each sample. This task is really painful and highly time-consuming. To adapt the system to the changes in the system and / or products, they should be renewed if necessary.

Then, colors which must be identified are subjective. Because of the impact of the wooden fiber to the perceived color of the wood, human operators can have different perceptions of the wood color. The output classes are gradual and non-disjointed too. For example, there are no strict bounds between a “red” wood and a “light red” wood.

Finally, there are other specific constraints to respect such as the real-time aspect of the production system or how easy the method is to set up.

So, the used classification method must take into account these constraints and present a low complexity of the recognition model.

The existing methods do not answer exactly to these specific constraints. Indeed, Neural Networks are popular machine learning algorithms that remain to be widely used for wood recognition problems. Their main advantages are their ease of use and their good behavior in term of classification rates. But their main drawback is that they need lots of training samples [15]. Other compilation methods such as Genetic Algorithm have the same drawback. The other most popular method in wood color recognition is k Nearest Neighbor. The k-NN is easily implemented as it does not require a training process. It is useful especially when there is a small dataset available. However, the major drawback of the Nearest Neighbor algorithms is that the computing time will increase according to the k number of neighbor used [15]. Moreover, the classification

is very dependent on the choice of these neighbours. So, the setting must be done by an expert in image processing.

We propose to use some tools provided by the Fuzzy Sets Theory [16], which seems to be well adapted for taking into account the above detailed constraints. Fuzzy Logic is able to provide non-separated output classes. While in the past fuzzy rule-based systems have been mainly applied to control problems, they have also been used recently in pattern classification tasks [17] [18] and they proved their ability to work with few learning data sets [19] [20]. That's why a classifier based on fuzzy linguistic rules seems to be more appropriate. Another advantage is the interpretability of such Rule Systems [17] [21] [22]. On the other hand, for instance, Neural Network are used as "black box" and there is no explicit link between the vocabularies used to define the output classes and the image features characterizing the color, it is very hard to interpret the decision taken by the classification method. The aim of this study is to obtain a vision sensor which delivers an answer in the "wood" vocabulary. This problem is often referred to as the "semantic gap", defined as "the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation" [23] Indeed, the wood expert will use words like "Red" or "Brown", whereas the vision expert will define colors with numerical values like {90, 35, 12} in RGB space for instance.

Moreover, the setting parameters must be comprehensible for non specialists too. In this way, it is judicious to be interested in the classification methods based on fuzzy set theory, because it allows the integration of the information expressed under linguistic form [24]. Finally, in highly specialized field like the timber industry, the experts make decisions whatever the work conditions. Thus, the vision system, which

must carry out the colorimetric control of the wooden boards, can be hatched as a decision system reproducing the human expert reasoning. In [25], D'Acquila recommends using inference engines which allows not only work from a representation space with n dimensions, but also takes into account all forms of uncertainty and imprecision.

This paper is organized as follows: Section 2 introduces the problems of the wood color recognition and the vision system used to do that. Section 3 details the proposed Fuzzy Rule Classifier by explaining the different steps of the method and its settings are explained in Section 4. Finally, Section 5 presents several comparisons with other classifiers and analyses the experimental results obtained from several University of California Irvine learning databases [26] and the industrial database corresponding to the application described in Section 2.

2. Industrial Vision Process

For about fifty years, the timber industry has been placed on a competitive market. In the order to get away from the industries which provide “bottom-of-the-range” products, some companies have given each other consequent means, like the use of vision systems in order to control and to enhance the production performances. As mentioned in the introduction, the context of this study is the classification of wooden colors.

2.1. Acquisition step

This recognition is carried out in real time on the industrial production line. These lines may reach speeds up to 400 meters of board length per minute. After the color identification step, done by the vision system, color information is sent to an

optimization step. Then each board is sent to a sorting line or to a cutting line. The cutting line aims to split the boards into uniformly colored piece of wood. The sorting line aims to group pieces of wood into specific classes, whose number and definition are given by the final customer. The boundary classes are very subjective in both cases.

The originality of the process concerns the color sorting which is only realized on the wooden board edges (board thickness). Indeed, the machining of handrails requires a uniform color in a large thickness (Fig. 1). To obtain this large thickness, three boards are glued by their face. So, the final product makes illusion of a product

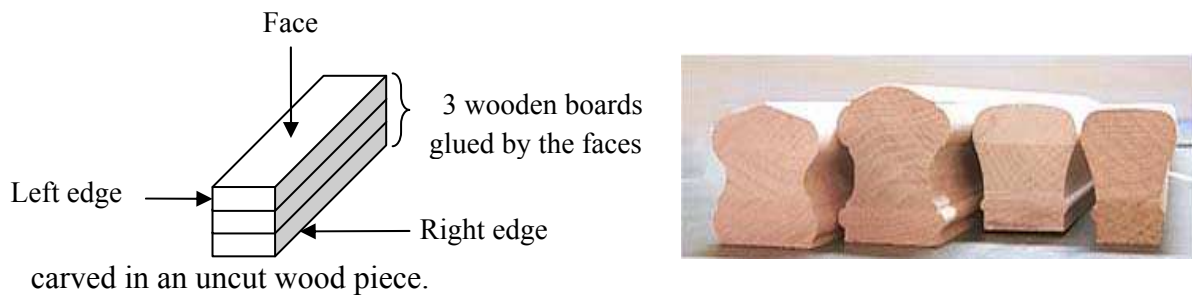


Fig. 1. Schematic representation of the final products and Handrail picture.

The entire vision system is detailed in Fig. 2. The acquired images are processed to obtain color descriptors (see Section 2.2). These features are then used by the classification stage (see Section 3) to provide the color label of the wooden board or a part of it if the color changes along the board.

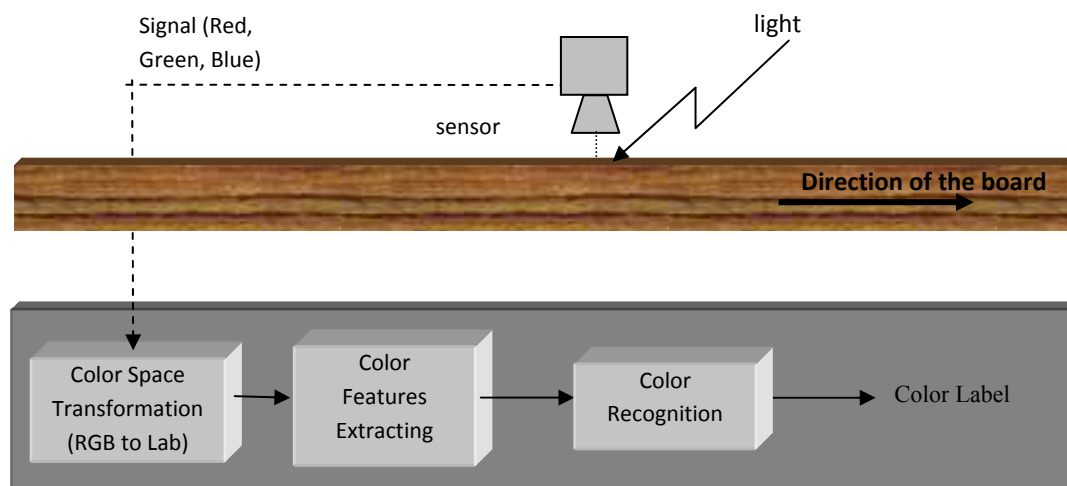


Fig. 2. Acquisition System and Processing Steps.

Fig.3 shows an image obtained with the acquisition system which is made up of one type of linear sensors: CCD color cameras. This CCD sensor provides the red, green and blue components of the signal. The signals are sampled at the rate of 1500 lines per second along y axis (Fig. 3). Each line is composed of 900 pixels (x axis on Fig.3). In the industrial case presented, the wooden boards are around three meters tall. With a longitudinal resolution equal to 1.5 millimeters per pixel, the images are made up of around 2000 lines. Thus, working in real-time, the data processing must be carried out under time constraints of around 1.5 seconds.

In this study, we work essentially with red oak because this wood species represents the most disadvantageous case. Indeed, for a particular wood hue, it is very hard to define with precision the customer color classes. The color variation can be very gradual (light red, medium red and dark red). Fig. 3 illustrates this gradual variability of the color.

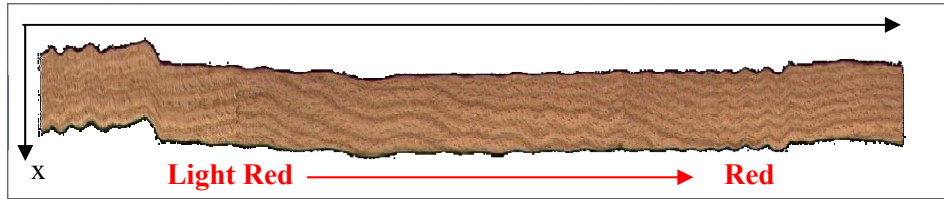


Fig. 3. Wooden board image with a gradual variation of the color from “Light Red” to “Red” class.

2.2. Characterization of the color

Two aspects are essential to characterizing color: the reference color space and the characteristic vector.

One of the most common color spaces denoted RGB, organizes the color information of an image into its red, green, and blue components. However, the International Commission on Illumination (CIE) does not recommend its use because the color components are not independent of one another [27]. Other popular color spaces include the Lab and HSV (Hue, Saturation, Value (intensity)) spaces. Many studies on color space selection have been conducted elsewhere, i.e. [28] [29]. After conducting several internal tests on various sets of wood samples, we decided to work in the Lab space because it provides the best color discrimination in term of recognition rates. We have also made this choice in relation to an objective criterion funded on Δ_{cielab} recommended distance [30]. This could perhaps be explained because this colorimetric reference space represents colors in the same order than humans do and the color class definitions are given by customers.

In the same way, it is necessary to characterize a color with a set of parameters which are extracted from the image. This set, called “characteristic vector” characterizes color in a simpler way. To answer to the real-time constraints imposed by the industrial production system, the characteristic vector must be easy to calculate. In order to choose

the best attributes to define the color, an extraction of 1D histogram features (statistical moments, entropy and homogeneity) from the color images is made. Then, a feature selection method, based on an adaptive algorithm of parameter space scaling using the Support Vector Machine (SVM) techniques [31], is applied. The suitability is determined from the calculation of the mean inertia of each feature according to a data set. The higher this value is, the more relevant the parameter is to distinguish the different classes. Fig. 4 illustrates the mean inertia of each parameter, calculated on the industrial training data set (see section 5.2 for details). This result shows that five features give a mean inertia higher than others.

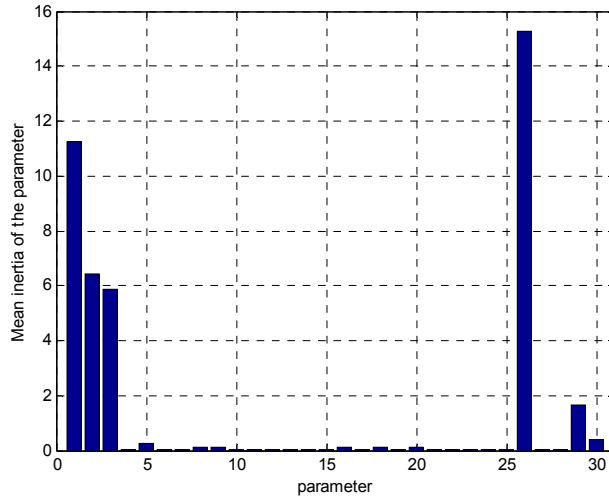


Fig. 4. Mean inertia of each parameter for the color discrimination.

The characteristic vector choice has been checked by taking into account the finality of the global system. Thus, the characteristic vector which provides the best recognition rate has been chosen [30]:

$$V_c = (m_L; m_a; m_b; hom_L; hom_h).$$

where m_x is the average of the component x , hom_x is the homogeneity of the component x : $hom_x = \sum h_i(x) / i$ with $h_i(x)$ the histogram of x .

The characteristic vector is composed with the components L , a and b from the Lab color reference space, and the component h (hue) calculated from the components a and b with the following equation: $h = \arctan(b/a)$.

3. The Fuzzy Reasoning Classifier

The different specifications allow the choice of a classification method based on the use of a linguistic model. The system integrating linguistic rule models are techniques often used in the computer vision field [32]. These methods are comprised of three parts: a set of “IF... THEN...” rules, a database, and an inference engine which allows to interpret the different rules and to provide the final classification result.

The Fuzzy Reasoning Classifier [33] is based on a fuzzy linguistic rule mechanism. It is well adapted to the presented industrial application. Indeed, it presents a very good and efficient generalization from a few sample sets and is able to provide gradual membership for output classes. Its satisfactory behavior will be shown in section 5.2 by several comparisons with other classifiers such as k Nearest Neighbor (K-NN), Neural Networks (NN) or Support Vector Machine (SVM).

This implemented algorithm, for the fuzzy recognition method, is a supervised learning mechanism divided into two stages (Fig.5). The Training stage is done off-line in order to set the recognition system and Generalization stage is used to classify color board on-line. The training part is composed of three steps: the Input Fuzzification of

the characteristic vector, the Fuzzy Rule generation from training data set and the Rule Adjustment which is the iterative part of the algorithm.

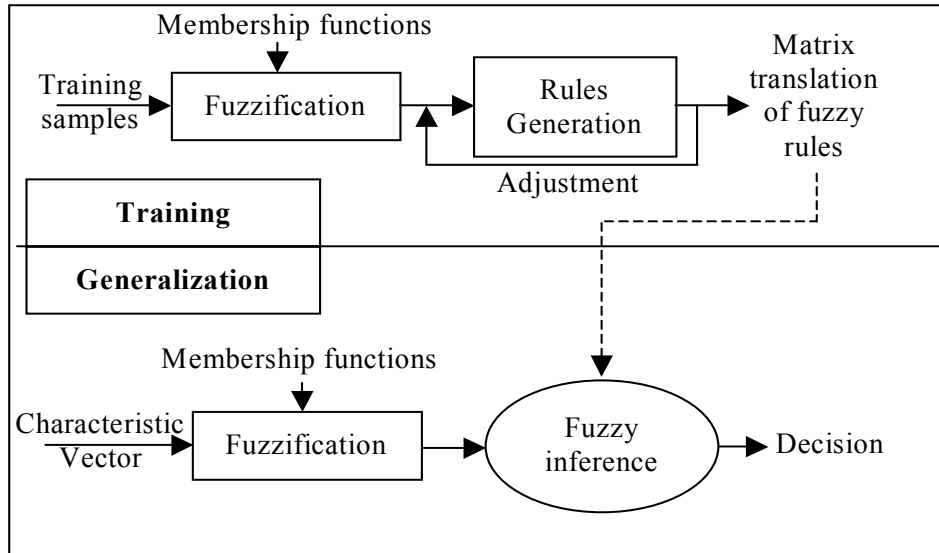


Fig. 5: Overall description of the fuzzy recognition system

These three steps are done off-line at the setting of the system. Then, these settings are used on-line. The membership functions defined in the fuzzification step are applied in the next step on the characteristic vector of “unknown” samples. The rule set which is automatically obtained from training without human checking, is used by the generalization step and the output class is determined by the rule of maximal answer. Fig.5 represents the different steps of the fuzzy recognition method.. It should be noted, that there is no defuzzification step in the proposed method because the exposed problem is quite different from the fuzzy control application. Symbolical outputs are needed, not numerical ones. Each output could be understood as a fuzzy singleton whose membership degree can be interpreted as the possibility of the sample to belong the considered class [34].

This output formalism contributes to reducing the difference existing between both industrial wood and image processing vocabularies [23]. The chosen formalism also improves the interpretability of the system.

3.1 Fuzzification of the characteristic vector

The fuzzification step aims to translate numerical variables into linguistic variables. A linguistic variable [35] is defined by a triple value (V, X, T_v) where:

- V is a variable (Luminance, Hue, etc.) defined on a set of reference X
- X is the universe of discourse (field of variation of V)
- T_v is the vocabulary chosen to describe in a symbolic way the values of V (weak, high, dark, light, etc.).

The set $T_v = \{A_1, A_2 \dots\}$, finite or infinite, contains normalized fuzzy subsets of X (also called Terms) which are usable to characterize V . Each fuzzy subset A_i is defined by the membership degree $\mu_{A_i}(x)$.

This fuzzification step defines the decomposition number of the considered variable to provide the fuzzy rule premises.

For example the membership function for variable L , called “Luminance”, is initialized with respect to the data analysis of the training sample set. The symbolic vocabulary then associated with the variable L is $T_L = \{\text{Weak, Medium, High}\}$. So, the linguistic variable “Luminance” is split into three Terms (or fuzzy sets) and this variable is characterized by a vector composed of three membership degrees: $[\mu_{\text{Weak}}(x), \mu_{\text{Medium}}(x), \mu_{\text{High}}(x)]^T$ as shown in Fig. 6.

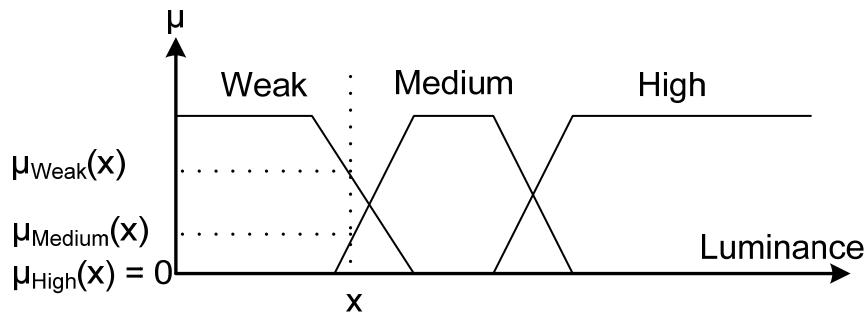


Fig. 6: Example of fuzzification of the parameter “Luminance” into 3 terms.

To summarize this step, one characteristic is represented by distributed terms in its definition field, called the universe of discourse, according to its useful and variable parts. The choice of the term number to be used to qualify a linguistic variable is one of the difficulties of this step. The others are to choose the shape of the membership function used to define these terms and to locate them on the universe of discourse. However, the industrial user, who is not an expert in pattern recognition, often chooses a regular distribution of terms, generally having more terms than are needed. Whenever the number of terms increases, so does the number of rules and, thus, the overall complexity of the entire system. Indeed, the rule number is given by Eq. (1):

$$\text{Numbers of Rules} = \prod_{v=1}^N \text{Card}(T_v) \quad (1)$$

with N the number of feature and $\text{Card}(T_v)$ the number of terms for variable V .

The choice of the fuzzification best adapted to the application problem is presented in section 4.

3.2 Fuzzy rule generation

This second step allows the defining of “IF... THEN...” fuzzy rules.

If two linguistic variables are considered for input (V_1, V_2) and one for output (Z_3), the general form of the associated fuzzy rule is [36]:

$$\mathbf{IF } V_1 \text{ is } A_i \mathbf{ AND } V_2 \text{ is } A_j \mathbf{ THEN } Z \text{ is } C_k \quad (2)$$

V_1, V_2 : input linguistic variables defined on X_1 and X_2 (intensity and surface for example)

Z_3 : output linguistic variable defined on Y (color name :”light Red” for instance)

A_i and $A_j \in Tv$ (chosen vocabulary: “weak” and “high” for example)

C_k : class of k^{th} color

Each rule describes the perceived defect, related to the system. Such rules can be classified into two categories: conjunctive rules and implicative rules. These two categories are regrouped, respectively. On the one hand, there are the possibility rules and the anti-gradual rules and, on the other hand, the certitude rules and the gradual rules [37]. The conjunctive rules are derived from the data analysis field where reasoning mechanisms are led by the data whereas implicative rules are most utilized in the cognitive sciences field where reasoning is led by knowledge [38].

For this application, conjunctive reasoning mechanisms have been logically selected. Each rule is activated in parallel and a disjunction operator combines the intermediate results. This inference mechanism gives an interpretation and semantics, which differ from mechanisms using implications [38]. In particular, it assures the consistency of the rule base [39]. If no information is processed, that is, the input space is not covered by the rule set; the output gives an “unknown defect” class. The two main models using these rules are the Larsen’s model and the Mamdani’s model [40]. The Sugeno’s model

[41] is not suitable in this case because the aim is not to achieve numerical output values.

There are many techniques to generate automatically linguistic rules. Alcalá et al. counts some methods [42]. The genetic algorithms are often used for this step [43-45] or the Decision Tree Method [46] [47]. Nevertheless, these techniques generally allow the fuzzification step and the rule generation step to be carried in the same way. Other methods are used too: the Wang and Mendel algorithm [48], an algorithm based on a weight average of the trained outputs [42], a mixed genetic algorithm based on a combination of the Wang and Mendel algorithm and an algorithm of fuzzy splitting of the discourse universes [42]. Table 1 summarizes a comparison of these different methods applied on the industrial training database (see Section 5.2).

Table 1
Recognition rates obtained from different methods of rule generation

Ishibushi's iterative algorithm [51]	Wang and Mendel's algorithm [48]	Weight average of the trained outputs [42]	Mixed genetic algorithm [42]
85.17%	81.02%	80.06%	83,73%

The results obtained with the algorithms presented by Alcalá et al. have two problems. The algorithm based on a weight average of the trained outputs is well adapted when the output classes are very distinct. The mixed genetic algorithm needs a consequent set of learning data to provide a good model and in our industrial context such a large set cannot be easily obtained.

So, the chosen classifier provides interesting results. It is based on Ishibuchi's algorithm which provides an automatic rule generation step [49] detailed below. Moreover, its inference mechanism follows the Larsen's model, which is better than the

Mamdani's model, because the Product is more adapted than the Minimum for the manipulation of several premises [50]. In fact, it allows non-linear splitting of variable input space. The iterative version of the Ishibushi's algorithm [51] [52] is used here. It allows the adjustment of the input space splitting by supporting the rule of having the maximum response. The entire algorithm is given in appendix.

In this algorithm the "AND" of Eq. (2) corresponds to the Cartesian product between V_1 and V_2 linguistic variables [40]. This operation is done with a T-Norm. A product is used in the algorithm (Eq. 3):

$$T(x_1, x_2) = \mu_A(x_1) * \mu_B(x_2) \quad (3)$$

The IF ... THEN implication is done through the Generalized Modus Ponens mechanism based on the use of Maximum/Product composition law. This inference follows Larsen's model [40], which uses a pseudo-implication operator represented by the product:

$$\text{IF_THEN} (V_1, V_2, Z) = T(T(x_1, x_2), y) \quad (4)$$

$$\beta_{Ck} = [\mu_A(x_1) * \mu_B(x_2)] * \mu_Z(y) \quad (5)$$

Eq. 4 corresponds to the expression Eq. 5 in the algorithm. Then, each rule gives a partial conclusion. β_{Ck} is aggregated to the others according to a fuzzy operator of disjunction. The disjunction operator is represented by the maximum operator according to Zadeh's case:

$$\beta_{CX} = \max \{ \beta_{C1}, \beta_{C2}, \dots, \beta_{CM} \} \quad (6)$$

where β_{CX} corresponds to the maximum membership degree given by the rule defined on $A_i \times A_j$.

Finally, CF_{ij} confident coefficient is calculated for all the M classes as follow:

$$CF_{ij} = (\beta_{CX} - \beta) / \sum_{T=1}^M \beta_{CT} \quad (7)$$

$$\text{where } \beta = \sum_{\substack{T=1 \\ CT \neq CX}}^M \beta_{CT} / (M - 1) \quad \text{and} \quad \beta_{CT} = \sum_{x \in CT} \mu_i(x_1) \times \mu_j(x_2)$$

CF_{ij} coefficient is associated to each rule R_{ij} , and is automatically adjusted in an iterative step.

3.3 Rule adjustment

The adjustment represents the iterative part of the algorithm. The following mechanism allows for adjusting the splitting of representative space according to the achieved results [51]:

- From the training patterns, the algorithm generates the first model.
- If the classification rate is below a ε threshold, defined by the user, the iterative part is performed to adjust this rate.

In fact, fuzzy rules are generated again by both injecting the training patterns and considering the new response of each rule and adjusting the CF_{ij} confident coefficient with the following equations.

When x is properly classified by the R_{ij} rule, the adjustment of the CF_{ij} confident coefficient is done by:

$$CF_{ij} = CF_{ij} + \eta_1(1 - CF_{ij}) \quad (8)$$

On the opposite side, when x is poorly classified by the R_{ij} rule, the adjustment of the CF confident coefficient is done by:

$$CF_{ij} = CF_{ij} - (\eta_2 \times CF_{ij}) \quad (9)$$

The rule set is automatically generated from learning data, without human checking. The CF_{ij} confident coefficient can be considered as a truth degree of a rule. An example of a generated rule set is given in Table 2. It shows the obtained matrix C (Fig. 4) whose first line could be interpreted as: “**IF** *Luminance L* is Weak **AND** *Chrominance a* is Low **AND** *Chrominance b* is Low **AND**... **THEN** Output Class is *Dark Brown* with *CF* confident coefficient equal to 0.851.

Table 2
Extract of Matrix C, representing the rule set provided by the generation step.

Premises					Output Classes	CF
L	a	b	hom _L	hom _h		
Weak	Low	Low	Low	Weak	DB	0.851
Weak	Low	Low	Low	Medium	DB	0.0004
...
Weak	Low	Medium	Medium	High	DR	1
...

The algorithm proposes an additional refining step. This step allows the improvement of the membership degree of the maximum membership class by modifying the slope of its membership function. This way is not studied here due to the gradual answer which needs to be kept because this vagueness improves the generalization capability of the classifier.

In the same way, recent works extend the algorithm by allowing the incorporation of weighted training samples [18]. In our case, the enhancement cannot be used because no information about training samples can be obtained.

3.4 “Unknown” sample classification:

There are two steps to classify an “unknown” sample from the rule set previously obtained:

- 1) to compute α_{CT} for each T class (T = 1 .. M):

$$\alpha_{CT} = \max(\mu_i^I(x_A) * \mu_j^J(xB) * CF_{ij}^{IJ}) \quad (10)$$

- 2) to allocate sample x to the X class such as:

$$\alpha_{CX} = \max\{\alpha_{C1}, \alpha_{C2}, \dots, \alpha_{CT}\} \quad (11)$$

Thus, each “unknown” wooden color sample will be classified with a membership degree α_{CT} to a symbolic color class T. This membership degree can be interpreted as a certainty degree corresponding to the possibility of the sample x to belong to the class T according to [34].

4 Fuzzy Rule Classifier Settings

The main settings of the Fuzzy Rule Classifier deals with the Fuzzification step. It has a great influence on the classification result. In this step, the number of Fuzzification Terms, the shape of the membership functions defining them and their position on the universe of discourse. In this section we present two ways in order to fuzzify the input variable before using by the rule set. Thus, we present a regularly distributed Fuzzification and an original automatic one. All the tests, presented in this section, are done with the industrial generalization data set (see Section 5.2).

4.1 Equally distributed fuzzification

This is the simplest fuzzification way which consists in splitting the variable universe of discourse into regular parts, depending on the term number.

The term number is empirically chosen from tests on input data in relation to the application field. Generally, this number is odd and small in order to limit the rule number (see Eq. (1)). The shape of the membership functions, describing terms have also to be fixed. Table 3 presents the results obtained from an equal-distributed fuzzification with the FRC. The different parameters of the characteristic vector are fuzzified with the same number of terms and with the same shape of membership function. These tests have been done with Matlab Fuzzy Toolbox where the membership functions have been represented with mathematical expression.

Table 3
Recognition rates obtained with an equal-distributed fuzzification for different shape of fuzzification curves on industrial generalization database.

Number of terms	Triangular Membership Functions	Trapezoidal Membership Functions	Gaussian Membership Functions	Triangular/Trapezoidal Membership Functions	Gaussian trapezoidal Membership Functions
2	45.30%	47.21%	41.95%	47.21%	44.98%
3	76.56%	77.03%	56.14%	77.35%	64.27%
5	81.82%	82.14%	78.47%	82.78%	79.11%
7	83.25%	83.89%	78.63%	84.37%	79.90%
9	78.47%	78.31%	73.21%	79.90%	74.48%
11	73.52%	73.84%	72.89%	75.28%	73.05%

For each shape of membership function, the best classification results are obtained by splitting the variables in seven terms. Decreasing rates appear when the splitting is more important. In this case, it is too precise to be representative of the classes to be identified. In relation with these tests, triangular/trapezoidal membership

function shape is chosen because they provide the highest recognition rate (84.37%). It should be noted that for the best rate, the rule number is equal to 16 807 (see Eq. 1).

In section 5, the presented results are obtained with this kind of curves. In section 5.1, the number of terms is fixed to 7. But the regular distributed fuzzification does not fit with the industrial case because it generates too many rules and low recognition rates.

Another way is to manually adapt the number and the position of these membership functions.

4.1 Adapted fuzzification

The adapted fuzzification needs expert knowledge. In this case, the choice of the term number and their position are determined from human interpretation of the data or from knowledge about the industrial context. We tried to model expert knowledge used to set up the Fuzzification step [53]. By carrying out one model for each expert (“wood” and “vision”), it is possible to obtain an adapted fuzzification which as described in Table 4 for the used characteristic vector.

Table 4
Number of terms chosen by Expert for an Adapted Fuzzification on industrial training data set.

Feature	m_L	m_a	m_b	hom_L	hom_h
Number of Terms	3	2	2	2	2

With this fuzzification, the recognition rate reaches **83.25%** on the generalization industrial database. By comparing this rate with the ones obtained with an equal-distributed fuzzification, a 1%-decreasing of the recognition rates is showed. Nevertheless, this fuzzification allows the reduction of the number of rules which is

equal to 48 according to Eq 1. The main advantage is that such a rule set becomes interpretable and adding or modifying a rule manually will be easier.

However, this method has the drawback of needing an expert to set up the FRC. So in the following section we propose an automatic method to fix the number and the position of the membership function of each variable.

4.2 Automatic fuzzification

The automatic fuzzification is directly linked to the analysis of the learning data set. In this case, only the numerical data are necessary in order to evaluate the better splitting of the discourse universe for each parameters of the characteristic vector. The advantage of such a fuzzification resides in the simplicity of the use. Indeed, the users of the classification systems, in an industrial context, are not experts, and thus prefer to quickly configure the recognition module. That is why they use often an equal-distributed fuzzification even if it is not adapted to the problem.

There are many techniques for carrying out the splitting of the discourse universes. They are usually based on clustering methods [54] [55] or on genetic algorithms [56-58]. Nevertheless, considering the genetic algorithm, it is very hard to obtain the right results in relation to the data sets used in our study. Indeed, like the neural networks, these techniques are known to be used with lots of learning points. Moreover, if the partition of the input variable space does not fit with the real data, the terms and the number of terms will be inappropriate.

That is why a fuzzification method based on the study of the output class typicality scores is used. Typicality measure $T(V)$ is calculated with the following equation from extern dissimilarity and intern likeness according to output classes [59].

$$T(x_a^u) = \frac{R \cdot D}{R \cdot D + (1 - R) \cdot (1 - D)} \quad (12)$$

$$R(x_a^u) = \frac{\sum_{i=1}^n \frac{1}{d(x_a^u, x_a^{f_i})}}{n} \quad (13)$$

$$D(x_a^u) = \frac{\sum_{i=1}^m 1 - d(x_a^u, x_a^{e_i})}{m} \quad (14)$$

where x_a^u is value of parameter a for sample x,

$x_a^{f_i}$ is value of parameter a for sample f belonging to the same class than x,

$x_a^{e_i}$ is value of feature a for sample e not belonging to the same class than x,

$d(x, y)$ is Euclidian distance,

n is the number of samples which belongs to the same class than sample x,

m is the number of samples which does not belong to the same class than sample x.

From Typicality measure $T(V)$, correlation (Corr) and cross-correlation (Xcorr) coefficients are calculated for each output classes. Then, from the ratio Corr/Xcorr, which characterizes the inter-classes similarity, the number of terms is determined. These terms are represented by Triangular/Trapezoidal Membership Functions which gives the best results (see Section 4.1). Their positions are obtained by computing the mean value of the samples belonging to the considered output classes [20]. Fig. 7 provides an example of automatic Fuzzification obtained with the proposed method and Table 5 gives the term number achieved for the entire characteristic vector of the application. This fuzzification has been obtained from the industrial training data set.

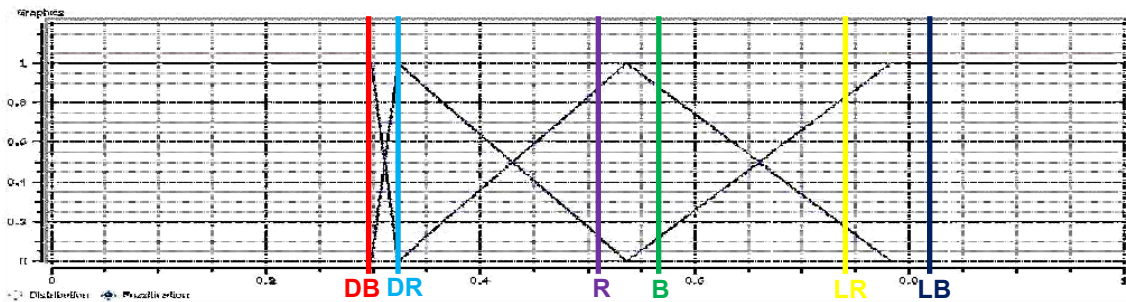


Fig. 7: Example of Automatic fuzzification of the “Luminance” parameter into 4 subsets – Position of the average of each class (DB, DR, R, B, LR, LB).

Table 5
Number of terms provided by the proposed Automatic Fuzzification method from the industrial training data set.

Feature	m_L	m_a	m_b	hom_L	hom_h
Number of Terms	4	3	5	3	4

The main interest takes place in the automatic adaptation of the fuzzification step which makes the tuning of the system easier. Table 6 illustrates the results obtained on the industrial generalization data set with the FRC and the above-quoted automatic fuzzification methods. It also gives the number of generated rules according to the retained term number.

Table 6
Recognition rates and rule number obtained with Automatic Fuzzification methods on the industrial generalization data set.

Clustering method (fuzzy c-means)		Genetic Algorithms		Typicality score analysis	
Recognition rate	Rule number	Recognition rate	Rule number	Recognition rate	Rule number
85.17%	1200	84.37%	960	85.65%	720

From all these tests concerning the fuzzification methods, two enhancements appear: one about the recognition rates and one about the rule number. Firstly, by

comparing the Automatic Fuzzification to the others, a 2%-improvement of the recognition rates is observed. Finally, we notice a decreasing of the rule number for the automatic fuzzification in comparison with an equal-distributed Fuzzification. However, the number of generated rules with an automatic fuzzification is higher than the number obtained in the Adapted case. Thus the rule base is harder to interpret, but it corresponds best to the expectations of the customers in terms of good classification and simplicity of F.R.C. use.

5. Experimental results

In this section, some tests are proposed with the aim of comparing or situating the efficiency of the proposed Fuzzy Rule Classifier. These comparisons are done in relation with usual methods used in pattern recognition or classification problems. So we applied the following classifier on some of the University of California - Irvine learning databases and on the industrial data set: Bayesian classifier (Bayes), Decision Tree Method (DTM: C4.5 algorithm) [47], k Nearest Neighbor algorithm (kNN) [47] or its Fuzzy version (FkNN) [60], Neural Networks (NN) [61], Genetic Algorithms (GA) [14] and Support Vector Machine (SVM) [62] which are the newer trends in machine learning algorithm which has been popular in many pattern recognition problems in recent years [63].

Some results concerning the application of this classifier on UCI database are given.

The aim of testing UCI databases is to situate the efficiency of FRC in relation to other reference classifiers [26].

Then, the results about color classification on wooden boards will be shown. They attempt to show the good capability of the proposed classifier for generalization from few training samples.

5.1 UCI benchmark experimentations

Before giving the classification results concerning the industrial context, it is important to compare the proposed F.RC. Method on the University of California - Irvine learning databases. The comparison is made with usual methods whose setting are similar to those used in Section 5.2 except for k-NN, Fk-NN and FRC. For k-NN and F-k-NN methods, k is chosen equal to 3 and for FRC setting, fuzzication is equally distributed in 2 or 3 terms.

Table 7 gives the recognition rates obtained with several UCI databases:

- UCI Iris benchmark: it is composed of 150 samples divided into 3 classes. For each sample, 4 different features are calculated.
- UCI Liver benchmark: it is composed of 345 samples divided into 2 classes. For each sample, 6 different features are calculated.
- UCI Diabetes benchmark: it is composed of 768 samples divided into 2 classes. For each sample, 8 different features are calculated.
- UCI Glass benchmark: it is composed of 214 samples divided into 6 classes. For each sample, 9 different features are calculated.

Table 7
Comparison of the classification for five UCI benchmarks.

UCI Databases	Bayes	3-NN	F-3-NN	NN	DTM	SVM	FRC
IRIS	96.70%	98.90%	99.1%	99.10%	92.02%	97.80%	97.33%
GLASS	-	72.00%	91.13%	91.13%	90.35%	91.13%	91.13%
DIABETES	76.50%	70.30%	80.88%	-	79.97%	80.88%	81.54%
LIVER	57.30%	62.90%	73.99%	73.99%	73.26%	73.99%	77.55%

For the Iris and Glass benchmark, the obtained results show very similar recognition rates for all classifiers except for Decision Tree Method. These tests demonstrate the good behavior of the proposed FRC method. It also gives the best results for two other UCI databases. Comparative results on these UCI databases with other classifiers are shown in [64].

In addition to the classification rates, the Cohen kappa coefficient is interesting to evaluate the performance of the FRC. This coefficient is used to evaluate the quality of the classifiers [65]. The coefficient kappa is computed with the confusion matrix obtained in a classification step (14):

$$\hat{K} = \frac{N \sum_{i=1}^r X_{ii} - \sum_{i=1}^r (X_{i+} * X_{+i})}{N^2 - \sum_{i=1}^r (X_{i+} * X_{+i})} \quad (14)$$

Where

- N is the total number of samples
- X_{ii} is the value in the row i and the column i
- r is the number of output classes
- X_{i+} is the sample number of the row i
- X_{+i} is the sample number of the column i.

This coefficient ranges from -1 to 1. The higher the result is, the better the classification is. In the UCI presented problems, the obtained results are summarized in table 8.

Table 8
Comparison of the Kappa coefficient for five UCI benchmarks.

UCI Databases	Bayes	3-NN	F-3-NN	NN	DTM	SVM	FRC
IRIS	0.8710	0.9240	0.9460	0.9460	0.8710	0.9240	0.9240
GLASS	0.2793	0.4653	0.5765	0.5665	0.5019	0.5665	0.5765
DIABETES	0.7540	0.5493	0.7805	0.2210	0.7805	0.7805	0.8129
LIVER	0.4290	0.6547	0.7116	0.7116	0.7116	0.7116	0.7990

According to the scale defined by Landis and Koch [66], the results presented in the table 8 confirm that the FRC is comparable to the other classifiers and even gives the best results in three cases. Nevertheless, such academic databases cannot really prove the generalization capacities of the proposed FRC method. That is why an evaluation of the generalization capability of the FRC is developed in the next section.

5.2 Industrial experimentations

We use an industrial data set composed of 943 samples distributed in 6 classes. This data set is composed with 84 “Dark Brown” samples, 176 “Brown”, 259 “Light Brown”, 54 “Dark Red”, 197 “Red” and 173 “Light Red”. To compare the different classifiers, some samples were randomly generated with a Gaussian white noise. Thus, for these tests, a data set composed of 5000 samples per class for the training step and 5000 samples per class for the generalization step (Test Extended Database) was used. Fig. 8 represents the evolution of the recognition rates during the generalization step functions of the number of learning samples per output class.

These results correspond to an average of the results obtained after some repetitions (100 random runs of the learning samples).

Fig. 8 shows the evolution of the generalization recognition rates in relation to the training sample number. An asymptote appears for each classifier:

- 71.5% for the Bayesian classifier;
- 76% for the k nearest neighbor algorithm (k=5);
- 80.2% for the Fuzzy k Nearest Neighbor algorithm (k=5);
- 83.4% for the Neural Networks (5 input neurons, 3 hidden layers of 20 neurons, 6 output neurons, 1 500 iterations, Levenberg-Marquardt Algorithm);
- 84.4% for the Fuzzy Rule Classifier ($\eta_1=0.4$, $\eta_2=0.004$, 500 iterations, $\varepsilon= 92\%$);
- 85.4% for a Genetic Algorithm (reproduction rate = 0.3, mutation probability=0.1, 1000 iterations) and
- 85.5% for the Support Vector Machine (Gaussian RBF Kernel, standard deviation=0.1, C=50, step of 5, tolerance 10^{-5}).

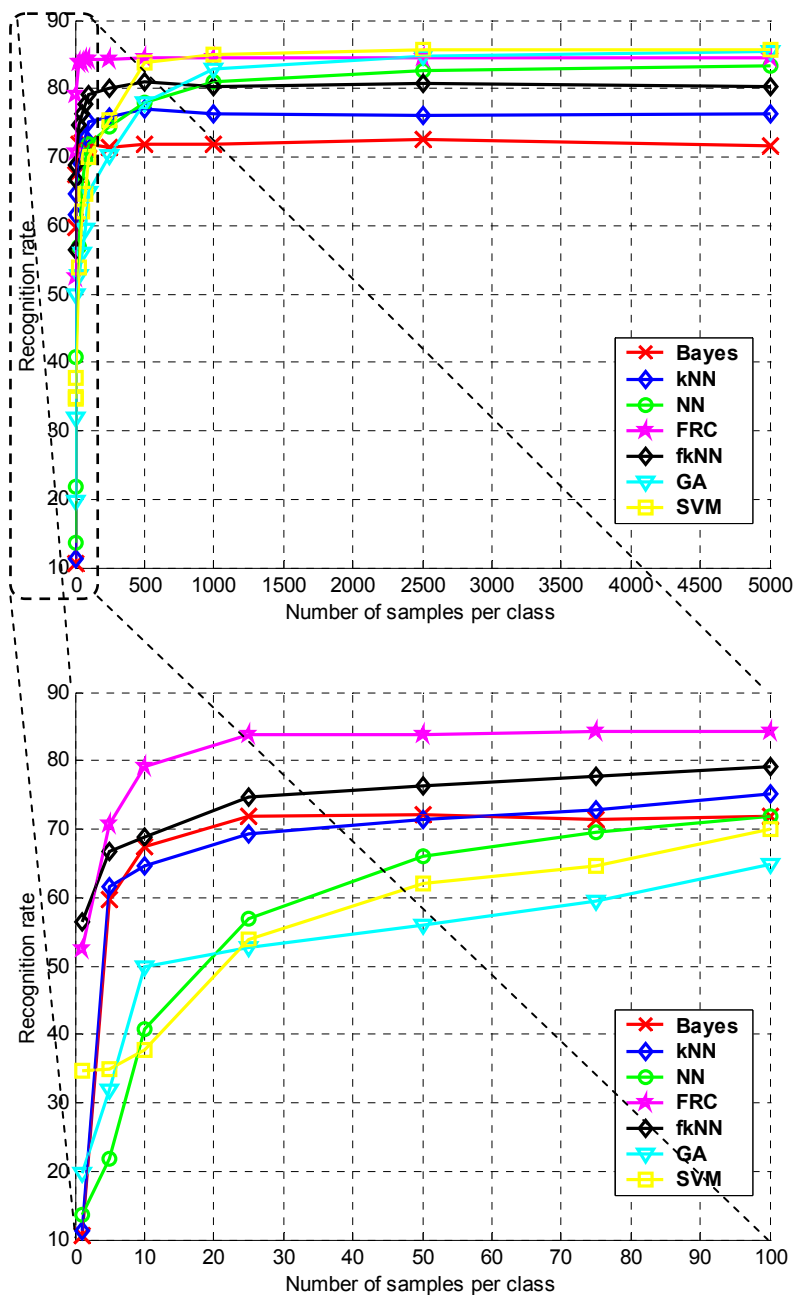


Fig. 8. Recognition rates functions of the sample number per class on the “test” extended database.

The FRC is one of the most effective classifiers. However, a difference exists concerning the convergence speed towards these maxima. Usually, the neural networks, the genetic algorithms and the SVM use lots of samples in the training step in order to give good results. For equivalent performances, these algorithms need more 100 as much (NN and GA) and 25 as much (SVM) of points so that the recognition rates are

comparable with that obtained with the FRC. Thus this aspect is very important, because it is not easy to have lots of training samples in an industrial environment. Moreover, in our applicative case, it is possible that some output classes are defined with few samples (about 10 or 20 samples). Thus, the methods which need lots of samples must be avoided.

The efficiency of the FRC method in generalization step has been shown and is strengthened in the industrial context. Fig. 9 illustrates the comparative results obtained with a “non-extended” training step. Indeed, in these cases, the training step was done with a dataset made of 25 randomly chosen samples of each class.

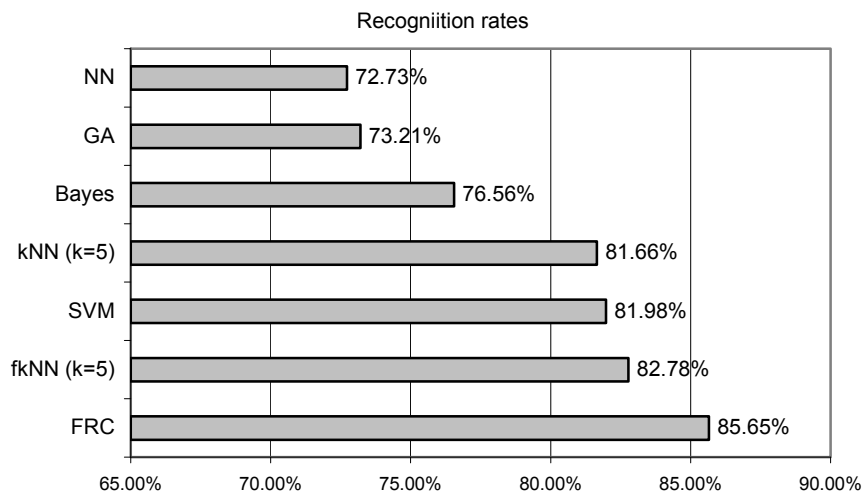


Fig. 9. Recognition rates obtained with the industrial generalization data set using 25 samples of each class for training.

These results show the same tendency. The lowest rates are obtained with two compilation methods which are effective with lots of training samples (NN and GA). The statistical methods, parametric or not, provide acceptable results. But FRC gives the best results (85.65%). This behavior is really interesting in our industrial case because of the low number of training samples needed. Finally, it should be noted that this data

set is given directly by the industrialist. This is no filtering and it is possible that there are many aberrant samples. It therefore shows a certain robustness and reliability of the proposed classification method. This method fits in well the time constraint of the system because it needs less than 1.30×10^{-4} s to recognize one sample (R.O.I.). It has been implemented in C/C++ with Microsoft XP and Intel core duo 3.00GHz CPU.

As UCI problems, the Kappa coefficient calculation allows to position our method. Table 9 presents these results which confirm the efficiency of the FRC classifier.

Table 9
Comparison of the Kappa coefficient for the industrial generalization data set.

NN	GA	Bayes	kNN	SVM	F-kNN	FRC
0.6578	0.6578	0.6991	0.7873	0.8003	0.8114	0.8291

6. Conclusions

In this paper, an original approach for a multi-class classification has been presented. Firstly, the efficiency of the FRC method has been presented through several comparisons on academic and industrial databases. Then, the interest of exploiting the fuzzy output of the FRC classifier has been shown in an industrial application case.

The proposed method uses Fuzzy Rules learned from small data sets. UCI benchmark results demonstrate that our method is comparable to other existing techniques like SVM, Genetic Algorithm, Neural Network, etc. The application of the method to an industrial problem (identification of wood board colors) has showed the efficiency, the robustness and the generalization capabilities of our proposition. In applicative cases, the setting of the method is very important to provide the best results, because the users

are not specialists of the fuzzy classification methods. The use of an automatic fuzzification method allows the improvement of the results in comparison with an equal-distributed fuzzification which is always employed by the users.

However, the global decision must take into account the decisions provided by two color sensors (right and left edges). The proposed Fuzzy Rules Classifier has been integrated in two fuzzy sensors. The results provided by these sensors are merged with a fuzzy operator detailed in [67]. The final recognition rates show an improvement of about 10% in comparison with a simple symbolic merging (according to the nominal single decision) [68].

One of the advantages of using fuzzy rules remains in the possibility of interpretation of the recognition model, but actually the size of the rule set does not allow it. So, the future work aims to reduce the complexity of the system in terms of model interpretability. Indeed, even if the results are interesting for an industrial production system, it should be important to be able to check the linguistic model from the expert knowledge [53]. Another way of investigation is to reduce the size of the characteristic vector by selecting suitable features. With this aim in view, an original method has been proposed [69]. It associates the FRC classifier with a suitable feature selection method, based on the Choquet integral.

Acknowledgment

The authors wish to thank their industrial partner who has provided the data sets and the knowledge concerning the framework.

Appendix

Outline of the Ishibushi rule generation algorithm.

A fuzzy rule R_{ij}^{IJ} for a two-dimensional classification problem can be written as follows:

Rule R_{ij}^{IJ} : If x_{1p} is A_i^I and x_{2p} is A_j^J then (x_{1p}, x_{2p}) belongs to Class C_{ij}^{IJ} with

$$CF = CF_{ij}^{IJ},$$

$$i=1, 2, \dots, I; I=1, 2, \dots, I_{max}; j=1, 2, \dots, J; J=1, 2, \dots, J_{max},$$

where C_{ij}^{IJ} is one of the M classes and CF_{ij}^{IJ} the certainty of the fuzzy rule R_{ij}^{IJ} .

(1) Calculate β_{CT} for $T = 1, 2, \dots, M$

$$\beta_{CT} = \sum_{x_p \in CT} \mu_i^I(x_{1p}) \times \mu_j^J(x_{2p})$$

(2) Find Class $X(CX)$ such that

$$\beta_{CX} = \max \{ \beta_{C1}, \beta_{C2}, \dots, \beta_{CM} \}$$

(3) CF_{ij}^{IJ} is determined as

$$CF_{ij}^{IJ} = (\beta_{CX} - \beta) / \sum_{T=1}^M \beta_{CT}$$

where

$$\beta = \sum_{\substack{T=1 \\ CT \neq CX}}^M \beta_{CT} / (M - 1)$$

References

- [1] T.H. Cho, R.W. Conners, P.A. Araman, A comparison of rule-based, K-nearest neighbour and neural net classifiers for automated industrial inspection. IEEE/ACM Int. Conf. on Developing and Managing Expert Systems Programs, October 1-2, Washington, DC, USA, 1991, pp 202-209.
- [2] D.L. Schmoltdt, R.M. Nelson, R.J. Ross, Ultrasonic defect detection in wooden pallets parts for quality sorting. Nondestructive Evaluation of Materials and Composites, December 3-5, Scottsdale, Arizona, USA, SPIE 2944 (1996) 285-295.
- [3] H. Kauppinen, H. Rautio, O. Silven, Non-segmenting defect detection and SOM based classification for surface inspection using color vision. Conference on Polarization and Color Techniques in Industrial Inspection, June 17-18, Munich, Germany, 1999, pp 270-280.
- [4] P.A. Estevez, C.A. Perez, E. Goles, Genetic input selection to a neural classifier for defect classification of radiata pine boards, Forest Products Journal 53 (2003) 87-94.
- [5] X. Bai, K. Wang, Research on classification of wood surface texture based on Markov random field. ICIEA 2007: Second IEEE Conference on Industrial Electronics and Applications, May 23-25, Harbin, China, 2007, pp. 664-668.
- [6] W. Pölzleitner, G. Schwingshagl, Real-time color-based texture analysis for sophisticated defect detection on wooden surfaces. 22th Conf on Intelligent robots and computer vision, October 24-27, Philadelphia, PA, USA, SPIE 5608 (2004) 54-69.
- [7] K. Wang, X. Bai, Research on classification of wood surface texture based on feature level data fusion. ICIEA 2007 Second IEEE Conference on Industrial Electronics and Applications, May 23-25, Harbin China, 2007, pp 659-663.
- [8] D.E. Kline, R.W. Conners, P.A. Araman, Technology to Sort Lumber by Color and Grain for Furniture Parts, Quality Lumber Drying in the Pacific Northwest, October 1-2 Seattle, United-States, 1999, pp 67-73.

- [9] T. Maenpaa, J. Viertola, M. Pietikainen, Optimising Colour and Texture Features for Real-time Visual Inspection. *Pattern Analysis Applications* 6 (2003) 169-175.
- [10] Q. Lu, A Real-Time System for Color Sorting Edge-Glued Panel Parts, Ph.D. Thesis, Virginia Polytechnic Institute and State University, 1997.
- [11] C. Daul, R. Rosch, B. Claus, Building a color classification system for textured and hue homogeneous surfaces: system calibration and algorithm, *Machine Vision and Applications* 12 (2000) 137-148.
- [12] A. Hanbury, Morphologie mathématique sur le cercle unité avec applications aux teintes et aux textures orientées. Phd thesis, Ecole Nationale Supérieure des Mines, Paris, 2002.
- [13] S. Srikanthswara, Q. Lu, W. King, T. Drayer, R. Connors, D.E. Kline, P. Araman, Real-time implementation of a color sorting system. *Machine Vision Applications, Architectures and Systems Integration*, October 15-16, Pittsburgh, PA, USA, 1997, pp. 170-179.
- [14] S. Sathyanath, F. Sahin, An AIS approach to a color image classification problem in a realtime industrial application, *International Conference on Systems, Man and Cybernetics* 4 (4) (2001) 2285-2290.
- [15] J. Y. Tou, Y. H. Tay, P. Y. Lau, Recent Trends in Texture Classification: a Review, *Symposium on Progress in Information & Communication Technology*, December 7-8, Kuala Lumpur, Malaysia, 2009, pp 64-68
- [16] L.A. Zadeh, Fuzzy sets, *Information and Controls* (1965) 338-353.
- [17] J. A. Roubos, M. Setnes J. Abonyi, Learning fuzzy classification rules from labeled data, *Information Sciences* 150 (2003) 77-93.
- [18] T. Nakashima, G. Schaefer, Y. Yokota, H. Ishibuchi, A weighted fuzzy classifier and its application to image processing tasks, *Fuzzy Sets and Systems* 158 (2007) 284 – 294.
- [19] F. Wang, L. Man, B. Wang, Y. Xiao, W. Pan, X. Lu, Fuzzy-based algorithm for color recognition of license plates, *Pattern Recognition Letters* 29 (2008) 1007-1020.

- [20] E. Schmitt, V. Bombardier, P. Charpentier, Self-Fuzzification Method according to Typicality Correlation for Classification on Tiny Data Sets, in: IEEE Conference on Fuzzy Systems, London, UK, 2007, pp. 1072-1077.
- [21] D. Nauck and R. Kruse, Nefclass-x – a soft computing tool to build readable fuzzy classifiers, *BT Technology Journal* 16 (3) (1998) 180–190.
- [22] D. Nauck, R. Kruse, Obtaining interpretable fuzzy classification rules from medical data, *Artificial Intelligence in Medicine* 16 (1999) 149–169.
- [23] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, Content-based image retrieval at the end of the early years, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22:12 (2000) 1349-1380.
- [24] D. Dubois, H. Prade, R.R. Yager, A Manifesto: Fuzzy Information Engineering, in: *Fuzzy Information Engineering: A Guided Tour of Applications (Hardcover)*, Wiley Publishers, 1996.
- [25] R.O. D’Acquila, C. Crespo, J.L. Mate, J. Pazos, An inference engine based on fuzzy logic for uncertain and imprecise expert reasoning, *Fuzzy Sets and Systems* 129 (2002) 187-202.
- [26] C. Blake, E. Keogh, C.J. Merz, UCI Repository of machine learning databases, University of California, Irvine.
- [27] International Commission on Illumination. Colorimetry, 2nd Edition, Publication CIE 15.2. <http://www.cie.co.at/>, 1986.
- [28] N.C. Burd, A.P. Dorey, Intelligent transducers, *Journal of Microcomputer Applications* 7 (1984) 87-97.
- [29] K. Leon, D. Mery, F. Pedreschi, F. Leon, Color measurement in L*a*b* units from RGB digital images, *Food research international* 39(2006) 1084-1091.
- [30] E. Schmitt, Contribution au Système d’Information d’un Produit Bois. Appariement automatique de pièces de bois selon des critères de couleur et de texture, Ph.D. Thesis, Henri Poincaré University, Nancy, France, 2007.

- [31] Y. Grandvalet, S. Canu, Adaptive Scaling for Feature Selection in SVMs, Neural Information Processing System MIT Press 15 (2002).
- [32] J.M. Keller, R. Krishnapuram, P.D. Gader, Y.S. Choi, Fuzzy Rule-Based Models in Computer Vision, in: Fuzzy Modelling: Paradigms and Practice, Kluwer Academic Publishers, Dordrecht, 1996, pp. 353-371.
- [33] E. Schmitt, C. Mazaud, V. Bombardier and P. Lhoste, A Fuzzy Reasoning Classification Method for Pattern Recognition, Proc. 15th Int. Conf. on Fuzzy Systems, FUZZIEEE'06, Vancouver, Canada, 2006, pp. 5998-6005.
- [34] D. Dubois, H. Prade, The three semantics of fuzzy sets , Fuzzy Sets and Systems 90 (1997) 141-150,.
- [35] H. J. Zimmermann, "Fuzzy set theory and its applications 4th Ed.", Kluwer Academic Ed., 2001.
- [36] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, Information Sciences (1975), Part-1 8 (3) 199-249, Part-2 8 (4) 301-357, Part-3 9 (1) 43-80.
- [37] D. Dubois and H. Prade, Fuzzy rules in knowledge-based systems - Modelling gradedness, uncertainty and preference, in: An introduction to fuzzy logic application in intelligent systems Kluwer, Dordrecht , 1992, pp. 45-68.
- [38] D. Dubois, H. Prade, What are fuzzy rules and how to use them, Fuzzy Sets and Systems 84 (1996) 169-185.
- [39] D. Dubois, H. Prade and L. Ughetto, Checking the coherence and redundancy of fuzzy knowledge bases, IEEE Trans. Fuzzy Systems 5 (1997) 398-417.
- [40] J.M. Mendel, Fuzzy logic systems for engineering: A tutorial, Proceedings of the IEEE 83 (1995) 345-377.
- [41] M. Sugeno, An introductory survey of fuzzy control, Information Sciences 36 (1985) 59-83.

- [42] R. Alcalá, J. Alcalá-Fdez, F. Herrera, J. Otero, Genetic learning of accurate and compact fuzzy rule based systems based on the 2-tuples linguistic representation, *International Journal of Approximate Reasoning* 44 (2007) 45-64.
- [43] O. Cordon, F. Herrera, P. Villar, Generating the knowledge base of a fuzzy rule-based system by the genetic learning of the data base, *IEEE Transaction of Fuzzy Systems* 9 (4) (2001) 667-674.
- [44] O. Cordon, M.J. Del Jesus, F. Herrera, A proposal on reasoning methods in fuzzy rule-based classification systems, *International Journal Of Approximate Reasoning* 20 (1999) 21-45.
- [45] H. Ishibuchi, T. Yamamoto, Effects of three-objective genetic rule selection on the generalization ability of fuzzy rule-based systems, 2nd Int. Conf. on Evolutionary Multi-Criterion Optimization – EMO'03, Faro, Portugal, 2003, pp. 608-622.
- [46] C. Marsala, "Fuzzy decision trees to help flexible querying", *Kybernetika* 36 (2000), 689-705.
- [47] D. Michie, Spiegelhalter, D.J., Taylor, C.C., *Machine Learning Neural and Statistical Classification*, Ellis Horwood, 1994.
- [48] L. Wang, J. Mendel, Generating fuzzy rules by learning from examples, *IEEE Transactions on Systems, Man, and Cybernetics* 22 (6) (1992) 1414-1427.
- [49] H. Ishibuchi, K. Nozaki, H. Tanaka, Distributed representation of fuzzy rules and its application to pattern classification, *Fuzzy Sets and Systems* 52 (1992) 21-32.
- [50] M.R. Berthold, Mixed fuzzy rule formation, *Fuzzy Sets and Systems* 32 (2003) 67-84.
- [51] K. Nozaki, H. Ishibuchi, H. Tanaka, A simple but powerful heuristic method for generating fuzzy rules from numerical data, *Fuzzy Sets and Systems* 86 (1997) 251-270.
- [52] H. Ishibuchi, T. Nakashima, Performance evaluation of fuzzy classifier systems for multi-dimensional pattern classification problems, *IEEE Trans. Systems, Man and Cybernetics - Part B: Cybernetics* 29 (1999) 601-618.

- [53] V. Bombardier, C. Mazaud, P. Lhoste, R. Vogrig, Contribution of fuzzy reasoning method to knowledge integration in a defect recognition system, *Computers in Industry* 58 (2007) 355-366.
- [54] T. Kempowsky, A. Subias, J. Aguilar-Martin, Process situation assessment: From a fuzzy partition to a finite state machine, *Engineering Applications of Artificial Intelligence* 19 (2006) 461-477.
- [55] F.A.T. De Carvalho, Fuzzy c-means clustering methods for symbolic interval data, *Pattern Recognition Letters* 28 (2007) 423-437.
- [56] H. Ishibuchi, K. Nozaki, N. Yamamoto, H. Tanaka, Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms, *Fuzzy Sets and Systems* 65 (1994) 237-253.
- [57] S.K. Oh, W. Pedrycz, H.S. Park, Hybrid identification in fuzzy neural-networks, *Fuzzy Sets and Systems* 138 (2003) 399-426.
- [58] O. Cordon, F. Gomide, F. Herrera, F. Hoffmann, L. Magdalena, Ten years of genetic fuzzy systems: current framework and new trends, *Fuzzy Sets and Systems* 141 (2004) 5-31.
- [59] J. Forest, M. Rifqi, B. Bouchon-Meunier, Class Segmentation to Improve Fuzzy Prototype Construction: Visualization and Characterization of Non Homogeneous Classes, in: *IEEE World Congress on Computational Intelligence, Vancouver, Canada, 2006*, pp 555-559.
- [60] J. M. Keller, M. R. Gray, and J. A. Givens, Jr., "A Fuzzy K-Nearest Neighbor Algorithm", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 15, No. 4, pp. 580-585, 1985.
- [61] D.T. Pham, S. Sagiroglu, Training multilayered perceptrons for pattern recognition: a comparative study of four training algorithms, *International Journal of Machine Tools & Manufacture* 41 (2001) 419-430.
- [62] P.Y. HAO, J.H. CHIANG, Y.K. TU, Hierarchically SVM classification based on support vector clustering method and its application to document categorization, *Expert Systems with Applications* 33 (2007) 627-635.
- [63] C. Chen, C. Chen and C. Chen, A Comparison of Texture Features Based on SVM and SOM, *ICPR* 2 (2006) 630-633.

- [64] M. Atif Tahir, A. Bouridane and F. Kurugollu, Simultaneous feature selection and feature weighting using Hybrid Tabu Search/K-nearest neighbor classifier, *Pattern Recognition Letters* 28 (2007) 438-446.
- [65] S. Garcia, A. Fernandez, J. Luengo, F. Herrera, A study of statistical techniques and performance measures for genetics-based machine learning: accuracy and interpretability, *Soft Computing* 13 (2009) 959-977.
- [66] J. R. Landis and G. G. Koch, The measurement of observer agreement for categorical data, *Biometrics* 33 (1977) 159-174.
- [67] O. Perez Oramas, Contribution to a methodology of knowledge integration for the image processing. Application to the contour detection by fuzzy linguistic rules (in french), Ph.D. Thesis, Henri Poincaré University, Nancy, France, 2000.
- [68] V. Bombardier, E. Schmitt, P. Charpentier, A fuzzy sensor for color matching vision system, *Measurement*, 42 (2) (2009) 189-201.
- [69] E. Schmitt, V. Bombardier, L. Wendling, Improving Fuzzy Rule Classifier by Extracting Suitable Features from Capacities with Respect to the Choquet Integral, *IEEE TSMC-B* 38 (5) (2008) 1995-1206.