Advances in Mathematical Sciences and Applications
Vol. 27, No. 1 (2018), pp. 65–80

GAKKOTOSHO
TOKYO JAPAN

A FUZZY-BASED APPROACH FOR THE LATIN MUSICAL GENRES INTELLIGENT CLASSIFICATION

Glaucia Maria Bressan *

Federal University of Technology of Parana, Mathematics Department
1640, Alberto Carazzai, 86300-000, Cornélio Procópio, Brazil
(E-mail: glauciabressan@utfpr.edu.br)

João Marcelo Abreu Bernardi

Federal University of Technology of Parana, Mechanical Engineering Department
1640, Alberto Carazzai, 86300-000, Cornélio Procópio, Brazil
(E-mail: joaomarcelo-14@hotmail.com)

and

Carlos Nascimento Silla Junior

Pontifical Catholic University of Parana, Graduate Program in Informatics
Curitiba, PR, Brazil
(E-mail: silla@ppgia.pucpr.br)

Abstract. This paper presents the construction of a fuzzy system for the automatic classification of Latin Musical genres. Although many techniques have already been proposed, no general solution to the problem exists, mainly due to the imprecise definition of musical genres. The Latin genres to be classified are extracted by [19] and they are called: tango, salsa, forró, axé, bachata, bolero, merengue, gaúcha, sertanejo and pagode. The system inputs features can be split into three groups: beat related (which includes the relative amplitudes and the beats per minute), timbre texture (the first five MFCCs) and pitch related (which includes the maximum periods and amplitudes of the pitch histograms). For each one of the ten musical genres, a fuzzy classification system is constructed for each of the three input groups. In the final step of this process, a membership percentage of an instance for each genre is obtained. The increasing number of musical genres, as well as their fusion and the influence that they receive and exert, motivates the use of fuzzy logic, since it is possible to consider uncertainties and the fuzziness among genres boundaries. The results of the classification are promising, as they suggest minor errors and classify data in a way close to the description of human decision criteria.

*corresponding author: glauciabressan@utfpr.edu.br

Communicated by Editors; Received September 26, 2017.
AMS Subject Classification: 68T10, 68T27, 68T05, 03E72.
Keywords: Modeling, Fuzzy systems, musical genres, neuro-fuzzy systems, intelligent classification.

1 Introduction

Musical genres classification has become a relevant problem, specially for the media industry. Providing properly selected music contents is fundamental for this industry, which manages huge music catalogues stored on distributed databases [6]. In this way, the classification in genres aims to group a music according to common characteristics. However, when these characteristics are mixed, sub-genres or styles from fusion are generated in an endless process, since each one of these productions is constantly influenced by other genres. Dividing music in genres is an attempt to classify each musical composition by considering objective criteria. The problem with automatic classification processes is that no clear definition and borders can be easily established.

The internet is destined to become the dominant medium for disseminating recorded multimedia content. As more music is made available via networks, the need for sophisticated methods to query and retrieve information from these musical databases increases [4]. Traditionally, the music classification has been performed as a manual process. This action requires a large amount of human effort and dedication. The correct association between musical genres and musical pieces is very important for several applications, for example, musical recommendation systems. Due to the increasing number of musical genres and their fusion, this traditional manual method has become obsolete. In addiction, human decision process can bring doubts and hesitations in the classification task.

Considering authors that applied fuzzy rules and classifiers to the musical genres classification problem, [8] present a preliminary attempt to apply Fuzzy Rule-Based System in cooperation with Evolutionary Algorithms to musical genres classification. One second of music is randomly extracted and the frequency spectrum is computed by means of the Fast Fourier Transform. Then, frequencies with higher energies are extracted and the relationship among them computed. The frequencies with higher energies -from first to fourth - are the inputs of the fuzzy system. Only two genres are considered: jazz and classical. According to the authors, this was the first time the fuzzy approach was applied to this kind of problem. After that, [6] developed a more systematic study. By applying the methodology on a larger training set, and by only using spectral information, the approach shows its capabilities for classifying previously unseen audio samples. The authors present a set of training-and-test experiments on a larger number of samples -considering different recording intervals and sampling rate, providing thus a series of results that show the usefulness of the approach - in this study, the same genres as before, jazz and classical, were considered. In their last paper, [7] present an ensemble of classifiers that uses a hybrid genetic fuzzy approach. By using a set of Fuzzy Rule Based Systems automatically tuned by means of a Genetic Algorithm, and structured in two layers, the system is capable of correctly classifying classical and jazz samples randomly chosen from a wide set of authors and styles.

Although some techniques have been performed in the literature, as it can be seen in [6, 7, 8, 9, 16, 17, 24], there is not any general solution to the musical genres automatic classification, specially due to the uncertainty among borders of genres. Also, only two genres are considered for classification. Besides, when an automatic classification is required, the extraction of input features and an efficient system, which must be able to process features, are necessary.

The method proposed in this paper can be used to automatically classify a larger number of genres, than the methods presented in the literature. Here, it is presented a case study considering Latin Music genres. The main goal of this work is to develop a fuzzy system in order to automatically classify Latin musical genres, using data from [19] and using input features from [20]. The Latin genres to be classified are: tango, salsa, forró, axé, bachata, bolero, merengue, gaúcha, sertanejo and pagode. The system inputs features can be split into three groups: beat related (which includes the relative amplitudes and the beats per minute), timbre texture (the first five MFCCs) and pitch related (which includes the maximum periods and amplitudes of the pitch histograms). For each one of the ten Latin musical genres, a fuzzy classification system is constructed for each of the three input groups. Neuro-fuzzy systems are considered to automatically generate the rule base and to fit the membership functions parameters. After the construction of a classification system for each one of the three input groups, a membership percentage is obtained using the arithmetic mean of the three outputs.

The contributions that can be highlighted, besides the automatic intelligent classification system, are: (i) the testing of the membership functions performances; (ii) the automatic generation of rules base and the membership functions parameters fit by using neuro-fuzzy systems; (iii) the proposal of a general solution to the musical genres classification, considering large number of genres and their fusions; (iv) the proposal of a recuperation system of musical information. For example, in a search process, users can select a musical genre that is mixed with another one.

In order to consider uncertainties and merges among borders of genres, in the automatic classification process, the use of fuzzy classification systems is motivated and promising, since fuzzy systems suggest minor errors and classify data in a way close to the description of human decision criteria, as described in [25] and [26].

This paper is organized as follows. Section 2 describes related works. Section 3 brings the Theoretical Reference, including Fuzzy Classification and Neuro-Fuzzy Systems concepts. After that, Section 4 presents results obtained from training neuro-fuzzy system and how the fuzzy classifier is designed. Results from classification are shown. Finally, Section 5 comments on the conclusions obtained from the results of classification.

2 Related Works

The automatic classification of songs into genres is one of the most popular tasks in the Music Information Retrieval research area and it was framed as a Pattern Recognition problem in the work of [22]. One of the additional contributions in the work of [22] was the GTZAN database that contains 30 audio clip from 1,000 songs that are classified into ten music genres. Despite some issues with this database [21] it was an important contribution to the field and it is still used in several studies up to this date.

Inspired by the work of [22] several researchers have address different issues in the task of music genre classification. The majority of the existing works focus on the development of one or more of the following aspects: (1) The creation of new music databases as music is a cultural phenomenon it is important to have different music databases with characteristics, instrumentation, cultural aspects, etc. Some of the existing examples of new databases are the RWC [10], USPOP [1], ISMIR 2004 and the Latin Music Database [19]. (2) The development of hand-crafted feature extraction algorithms such as Rhythm Histograms and Statistician Spectrum Descriptors [12]; and (3) The development of new approaches to classify music genres. To a more in depth discussion about the existing approaches the interested reader is referred to [13, 2]. A recent trend in the field is the use of deep learning techniques to automatically learn the feature representation from the data [14, 18]. Furthermore, it is possible to combine hand-crafted features with automatically learned ones to achieve better classification results [3].

Within the technical advances in the field, most of the work in the area still deals with the music genre classification problem as a single label classification problem. That is, each song have exactly one genre label that is belongs to. However, this can be a limited view as composers might want to make songs that draw some aspects from different styles. For this reason there are some works that deal with the music genre classification problem as a multi-label classification problem, where each song can be associated with a set of music genres [13].

Despite dealing with the classification of music genres as a single or a multi-label problem, the existing works are still focusing on the classification of music genres using the classic logic. A particular song either belongs or not to a given music genre. That is why most of the real world music information retrieval systems deals with music genre taxonomies for browsing songs, or use tags. However, one limitation that has not been addressed yet (to the best of the authors knowledge) is how to allow the users to search music genre databases with thresholds for the different music genres. For example, if a user is interested in rock music, that might have some classical aspect, current systems might be unable to provide the user with a set of relevant songs to his query.

One way to go towards that type of system is to use fuzzy-based system, that uses fuzzy logic and is able to measure how much a particular song belong to each of the different music genres in a music database. One of the limitations of employing a fuzzy-based system is that is necessary to model the rules within the system. As the boundaries between different music genres are fuzzy, in this work we also employ a neuro-fuzzy system to automatically learn the rules of the fuzzy-system.

3 Theoretical Reference

The main advantage of applying a fuzzy classification system is its ability to approach uncertainties in order to generate good classifiers. For this reason, it is a promising tool for the musical genres classification, due to inaccuracies among borders of genres.

3.1 Fuzzy Classification Systems

A fuzzy classification system is basically formed by the definition of the membership functions parameters, the input and output variables, and a base rule (called 'if-then' rules). Firstly, we have to define a *fuzzy set* [25].

Definition 1. A fuzzy set is characterized by a membership function mapping the elements of a domain, space or universe of discourse $\mathbf{X}$ to the unit interval [0,1], that is,

A: $\mathbf{X} \rightarrow [0,1]$.

Thus, a fuzzy set A in $\mathbf{X}$ may be represented as a set of ordered pairs of a generic element $x \in \mathbf{X}$ and its grade of membership: $A = \{(A(x)/x) | x \in \mathbf{X}\}$.

A membership function is defined by [15] as a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. The input space is sometimes referred to as the universe of discourse.

In principle any function of the form A: $\mathbf{X} \rightarrow [0,1]$ describes a membership function associated with a fuzzy set A that depends not only on the concept to be represented, but also on the context in which it is used [15]. Let m be a modal value, a and b the lower and upper bounds, respectively, for nonzero values of $A(x)$, the most common membership function models are (1), (2) and (3).

Triangular Functions

$$A(x) = \begin{cases} 0 & \text{se } x \leq a \\ \frac{x-a}{m-a} & \text{se } x \in [a, m] \\ \frac{b-x}{b-m} & \text{se } x \in [m, b] \\ 0 & \text{se } x \geq b \end{cases} \tag{1}$$

Trapezoidal Functions

$$A(x) = \begin{cases} 0 & \text{se } x < a \\ \frac{x-a}{m-a} & \text{se } x \in [a, m] \\ 1 & \text{se } x \in [m, n] \\ \frac{b-x}{b-n} & \text{se } x \in [n, b] \\ 0 & \text{se } x > b \end{cases} \tag{2}$$

Gaussian Functions

$$A(x) = e^{-k(x-m)^2}; k > 0 \tag{3}$$

The *triangular* membership function is a collection of three points forming a triangle. The *trapezoidal* has a flat top and the *Gaussian* membership function is built on the Gaussian distribution curve. The base rule, on the other hand, consists of conditional statements that comprise fuzzy logic. In general, the input to an if-then rule is the current value for the input variable and the output is an entire fuzzy set. The elements that structure a Fuzzy System can be seen in Figure 1. Fuzzy inference is the process of mapping from a given input to an output using fuzzy logic. The mapping then provides a basis from which decisions can be made. According to [15], the inference process involves 5 steps:

- *Fuzzify Inputs.* To take the inputs and determine the degree to which they belong to each of the appropriate fuzzy sets via membership functions. The input is always a crisp numerical value limited to the universe of discourse of the input variable and the output is a fuzzy degree of membership (the interval between 0 and 1). Once the inputs have been fuzzified, the degree to which each part of the antecedent has been satisfied for each rule is known.

- *Fuzzy Operator*. If the antecedent of a given rule has more than one part, the fuzzy operator is applied to obtain one number that represents the result of the antecedent for that rule. This number will then be applied to the output function. The input to the fuzzy operator is two or more membership values from fuzzified input variables. The output is a single truth value. Basically, AND (minimum) and OR (maximum) methods.

- *Implication Method*. It is defined as the shaping of the consequent (a fuzzy set) based on the antecedent (a single number). The input for the implication process is a single number given by the antecedent, and the output is a fuzzy set. Implication occurs for each rule. Two built-in methods are supported, and they are the same functions that are used by the AND method: minimum, which truncates the output fuzzy set, and product, which scales the output fuzzy set.

  These three steps above are the three steps of the if-then rule process.

- *Aggregation Method*. It is the process by which the fuzzy sets that represent the outputs of each rule are combined into a single fuzzy set. Aggregation only occurs once for each output variable. The output of this process is one fuzzy set for each output variable. Three built-in methods can be supported: max (maximum), probor (probabilistic or), and sum (simply the sum of each output set of the rule).

- *Defuzzify*. The input for the defuzzification process is a fuzzy set (the aggregate output fuzzy set) and the output is a single number. The most popular defuzzification method is the centroid calculation, which returns the center of area under the curve.

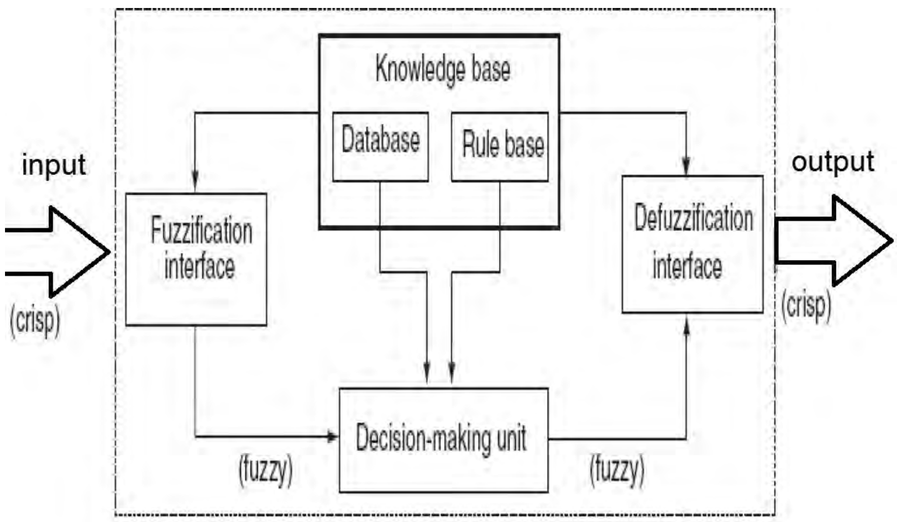

Figure 1: Inference Fuzzy System.

The classification system inputs for the problem addressed in this paper are attributes (inputs) selected by [20]. These attributes must be modeled by membership functions whose parameters and types (triangular, trapezoidal and Gaussian) are exhaustively tested. The membership functions parameters and the rules base are generated automatically by a Neuro-fuzzy System.

3.2 Neuro-fuzzy Systems

Since the creation of if-then rules and the fit of membership functions parameters are a laborious process, the neuro-fuzzy systems can automate the generation of these elements of the fuzzy system. The idea is to provide a method for the fuzzy modeling procedure to *learn* information about a data set, in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given imput/output data. This learning method works similarly to that of neural networks. The parameters associated with the membership functions will change through the learning process. The computation of these parameters (or their adjustment) is facilitated by a gradient vector, which provides a measure of how well the fuzzy inference system is modeling the input/output data for a given set of parameters. Once the gradient vector is obtained, any of several optimization routines could be applied in order to adjust the parameters so as to reduce some error measure. In this paper is used the hybrid method: a combination of least square estimation and backpropagation for membership function parameter estimation.

The Neuro-fuzzy system structure can be seen in Figure 2, which shows the layers of Neuro-fuzzy network, which are: 1 - crisp inputs; 2 - fuzzify inputs; 3 - rules base (linguistic); 4 - consequents and 5 - defuzzify.

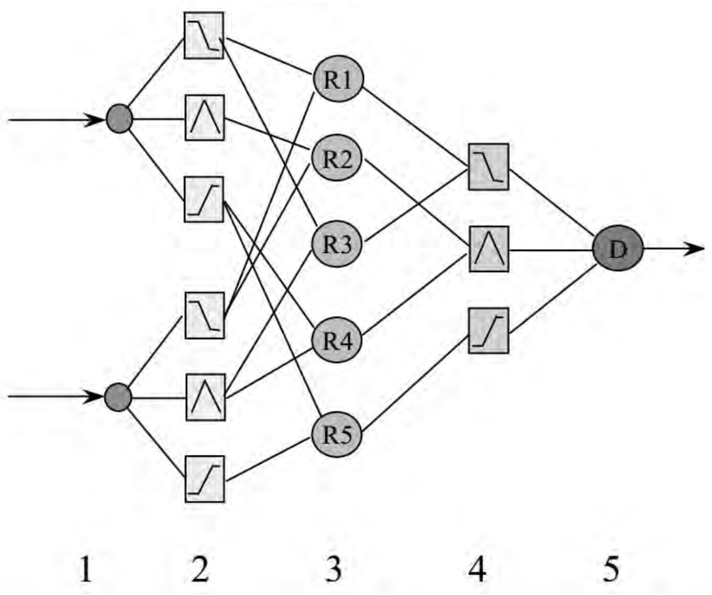

Figure 2: Neuro-fuzzy System Structure.

In the training data, each row is a desired input/output pair of the target system to be modeled; it starts with an input vector and is followed by an output value. Therefore the number of rows is equal to the number of training data pairs, and the number of columns is equal to the number of inputs plus one. Training option is a vector that specifies the stopping criteria: the training process stops if the designated epoch number is reached or the error goal is achieved. The checking data is used for testing the generalization capability of the fuzzy inference system at each epoch. The checking data has the same format as that of the training data, and its elements are usually distinct from those ones. This cross-validation gives an unbiased estimate of the minimal error measure that can be achieved in the training. This neuro-fuzzy system is trained using the Sugeno inference system, and this one must be of unity weight for each rule and linear or constant output membership functions. This adaptive neuro-fuzzy inference system can be seen with more details in [11].

4 Results

This section presents results obtained from the training neuro-fuzzy system - which corresponds to membership functions parameters and if-then rules - and from the fuzzy classification system.

4.1 Proposed Fuzzy Classifier

For each one of ten musical genres, a fuzzy system is constructed, as shown in Figure 3, for each of the three input groups: beat related (which includes the relative amplitudes and the beats per minute), timbral texture (the first five MFCCs) and pitch related (which includes the maximum periods and amplitudes of the pitch histograms). The features are extracted from [20]. The input database is a numeric (crisp) matrix, in which each column corresponds to a feature. The ten considered Latin genres to be classified are extracted by [19] and they are: tango, salsa, forró, axé, bachata, bolero, merengue, gaúcha, sertanejo and pagode.

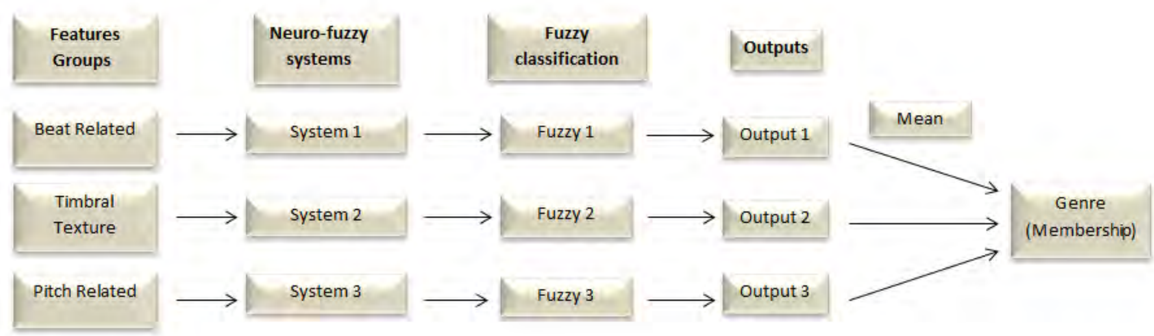

Figure 3: Fuzzy classification system structure for each musical genre.

In this paper, we used a set of the available music in *Latin Music Database*, from [19], a 3.000 musical recordings database from the ten different genres described above. This database is composed of 300 instances of each music genre, of which 10% are separated for testing. The MARSYAS framework was employed for feature extraction [23]. The features employed in this paper comprise short-time Fourier transform, Mel frequency cepstral coefficients (MFCC), beat and pitch related features, inter-onset interval histogram coefficients, rhythm histograms and statistical spectrum descriptors. More details can be seen in [20]. The features can be split into three groups: beat related (features 1-6), include the relative amplitudes and the beats per minute; timbral texture (features 7-11), the first five MFCCs; and pitch related features (features 12-16), include the maximum periods and amplitudes of the pitch peaks in the pitch histograms. Each row of the input matrix is called 'instance', it contains the numerical information about the features and it corresponds to one genre. The final feature vector is outlined in Table 1.

Table 1: Description of the feature vector

Feature	Description
1	Relative amplitude of the first histogram peak
2	Relative amplitude of the second histogram peak
3	Ratio between the amplitudes of the second peak and the first peak
4	Period of the first peak in bpm
5	Period of the second peak in bpm
6	Overall histogram sum (beat strength)
7	First MFCC mean
8	Second MFCC mean
9	Third MFCC mean
10	Fourth MFCC mean
11	Fifth MFCC mean
12	The overall sum of the histogram (pitch strength)
13	Period of the maximum peak of the unfolded histogram
14	Range of the maximum peak of the folded histogram
15	Period of the maximum peak of the folded histogram
16	Pitch interval between the two most prominent peaks of the folded histogram

4.2 Neuro-fuzzy System Structure

For each one of the ten Latin musical genres, 3 neuro-fuzzy systems are created: one system for each of the 3 features groups. Figure 4 shows the neuro-fuzzy structure for the "beat-related features" group, which is formed by features 1-6 from Table 1. In Figure 4, "input" represents the 6 beat-related features, according to Table 1. Each feature is partitioned in 3 classes (low, average and high) according to membership functions trained by the neuro-fuzzy system. These 3 classes are represented by "inputmf". After that, "rule" represents the if-then rule base, generated from classes combination. For this features group, $3^6 = 729$ rules are generated. The output of each rule is represented by "outputmf". Finally, "output" is the system output. In the neuro-fuzzy system we used 3 epochs to optimize the neuro-fuzzy and the hybrid method, described in Section 2.2, as the optimization process.

Thus, the extraction of knowledge is accomplished in fuzzy rules format. The crisp input features are partitioned in low, average and high classes, whose possible combinations generate if-then rules , which are connected by the logic operator "and".

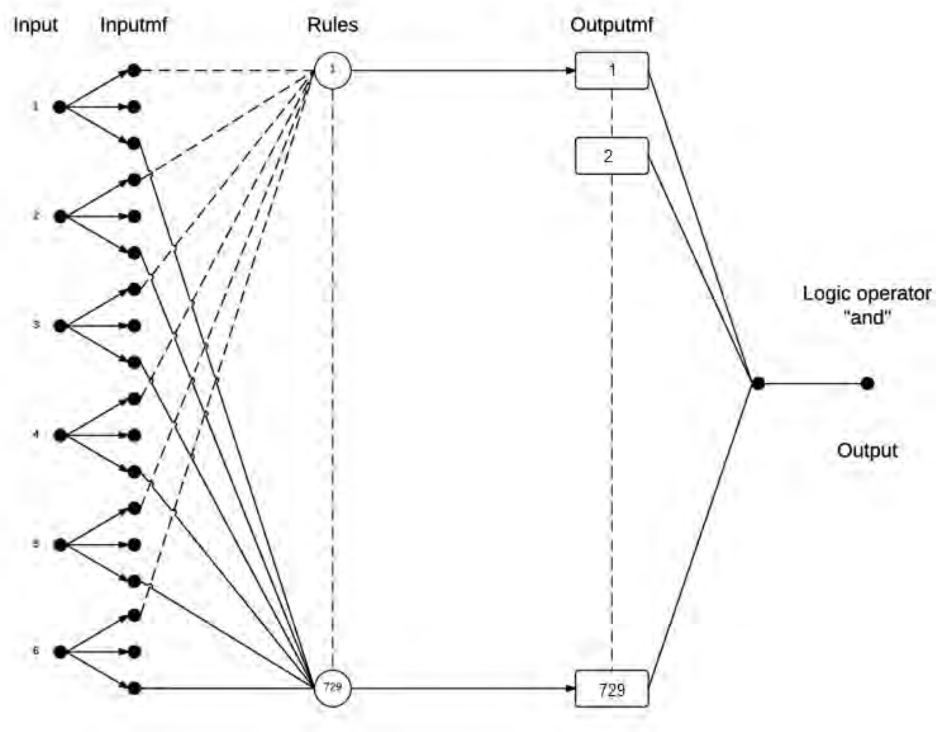

Figure 4: Neuro-fuzzy structure for "beat-related features" group.

4.3 Membership Functions

The input membership functions are modeled by the neuro-fuzzy system, as well as the numeric parameters fit. The shape of the membership functions employed in this problem is the Gaussian, since it presents the least significant error in relation to triangular and trapezoidal functions. In addiction, the Gaussian function also presents a gradual pass between the borders of classes. Since we have a classification system, as shown in Figure 3, for each one of the 10 Latin musical genres, we may consider as an example, the genre *tango*. The first group of features (1-6 features) is composed by 6 features. So, there are 6 membership functions, which are partitioned into 3 classes: low, average and high, according to the parameters specified in Table 2.

Table 2: Membership functions parameters: 1-6 features (beat-related features) for the *tango* genre.

Feature	Parameters		
	low	average	high
1	[-0.0003267 0.07694]	[0.07918 0.1891]	[0.07558 0.3793]
2	[0.007731 0.08286]	[0.07781 0.2049]	[0.07895 0.3988]
3	[0.1525 0.1783]	[0.1797 0.5761]	[0.1796 0.9999]
4	[5.143e-005 0.2123]	[0.2101 0.4998]	[0.2122 1]
5	[-2.038e-005 0.2121]	[0.2109 0.5]	[0.212 0.997]
6	[5.149e-005 0.2122]	[0.2114 0.5008]	[0.2107 1.001]
7	[0.01915 0.2083]	[0.2083 0.5096]	[0.2083 1]
8	[0.1635 0.1209]	[0.1209 0.4481]	[0.1209 0.7328]
9	[0.3744 0.1328]	[0.1328 0.6872]	[0.1328 1]
10	[0.003343 0.2116]	[0.2116 0.5017]	[0.212 1]
11	[0.08663 0.1939]	[0.1939 0.5433]	[0.1939 1]
12	[0.007771 0.0444]	[0.0425 0.11]	[0.0352 0.2148]
13	[-6.559e-005 0.1966]	[0.1967 0.4636]	[0.1966 0.927]
14	[0.01734 0.03879]	[0.03997 0.1104]	[0.03628 0.205]
15	[-0.0001094 0.2121]	[0.2115 0.4996]	[0.2122 1]
16	[-7.198e-005 0.1515]	[0.1502 0.3578]	[0.1512 0.7145]

The output membership functions are modeled as constant functions, since they are related to the ten Latin musical genres considered; in this example, the genre *tango*. The fuzzy inference system is Sugeno type [15]. Figure 5 illustrates the behavior membership functions, which are related to the System 1 in Figure 3, that is, the 6 features of the first group (beat-related features) of the *tango* genre.

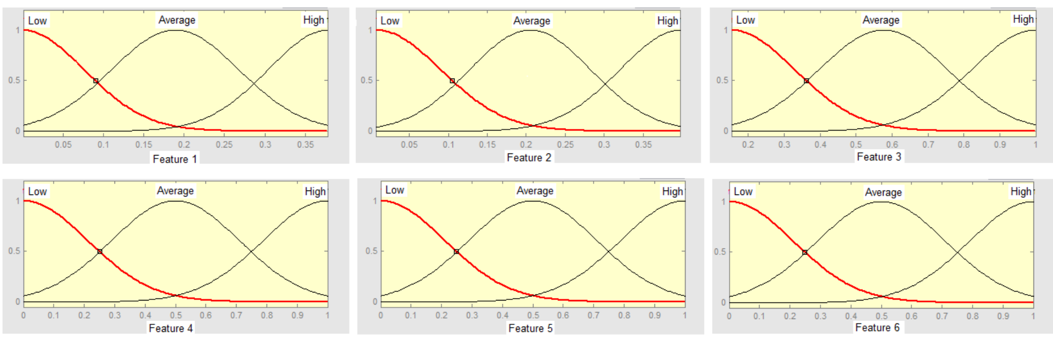

Figure 5: Membership functions related to the 6 features of beat-related features of the *tango* genre.

4.4 Rules Base

The 'if-then' rules base is, thus, generated from the neuro-fuzzy system. The inputs are logically combined applying the 'and' operator, which indicates that the inputs occurs simultaneously. Since classification system is composed of three input groups, whose data are divided into 3 classes, for each one of the 10 genres, the 6 features of the first group (beat-related features, 1-6 from Table 1) generated $3^6 = 729$ rules; the 5 features of the second group (timbral texture features, 7-11 from Table 1) generated $3^5 = 243$ rules and the last 5 features of the third group (the first five MFCCs and pitch related features, 12-16 from Table 1) also generate 243 rules from the features combination. For example, the rule number 1 of System 1 (see Figure 3) for the output *tango* is described as:

IF feature 1 is LOW, AND feature 2 is LOW, AND feature 3 is LOW, AND feature 4 is LOW, AND feature 5 is LOW, AND feature 6 is LOW, THEN genre is tango

After the rules base is generated and the membership functions parameters are fitted by the neuro-fuzzy system, these data are inserted into the fuzzy classification system in order to automatically classify the Latin musical genres considered, which are from [19].

4.5 Fuzzy Classification

As described in Section 3.1, from 3.000 musical recordings database composed of 300 instances of each musical genre considered in the paper [19], 10% are separated for testing, which correspond to 30 instances of each musical genre. The 30 test instances related to the *tango* genre, for example, are input to the Fuzzy Classification System structured for the *tango* genre, as illustrated in Figure 3, and then, an arithmetic mean is obtained from the three output systems. Similarly, other test instances, related to the other musical genres, are input to the corresponding Fuzzy Classification System and the final membership percentage is obtained. This process checks whether the output for the correct genre is high enough. The 30 test instances related to the each genre are input to the Fuzzy Classification System structured for the 9 others genres in order to check whether the outputs of the others are low enough. According to the membership function parameters, whose range is [0 1], for each test instance, if the arithmetic mean is between 0.995 and 1, then the system has classified the genre correctly. Otherwise, it has incorrectly classified. Table 3 describes the accuracy of each one of the 10 musical genres.

Table 3: Fuzzy classification accuracy for testing instances.

Genre	Correctly Classified	Incorrectly Classified	Total	Accuracy
Tango	28	2	30	93.3%
Salsa	29	1	30	96.7%
Forró	26	4	30	86.7%
Axé	29	1	30	96.7%
Bachata	26	4	30	86.7%
Bolero	26	4	30	86.7%
Merengue	26	4	30	86.7%
Gaúcha	27	3	30	90.0%
Sertanejo	26	4	30	86.7%
Pagode	29	1	30	96.7%

In order to compare the fuzzy classification with the traditional classifier called *k-nearest neighbor (k-NN)*, Table 4 shows the accuracy of each one of the 10 musical genres using the k-NN classifier. This algorithm assigns an input to a genre without considering uncertainties and fuzziness among genres boundaries. According to [5], in k-NN classification, the output is a class membership. An input is classified by a majority vote of its neighbors, with the input being assigned to the class most common among its k nearest neighbors Using $k = 1$, then the input is simply assigned to the class of that single nearest neighbor.

Table 4: k-NN classification accuracy for testing instances.

Genre	Correctly Classified	Incorrectly Classified	Total	Accuracy
Tango	26	4	30	86.7%
Salsa	28	2	30	93.3%
Forró	22	8	30	73.3%
Axé	24	6	30	80.0%
Bachata	24	6	30	80.0%
Bolero	23	7	30	76.7%
Merengue	23	7	30	76.7%
Gaúcha	27	3	30	90.0%
Sertanejo	25	5	30	83.3%
Pagode	28	2	30	93.3%

Considering the 300 total testing instances, the global success percentage of fuzzy classification is 90.7%, since 272 instances have been correctly classified. By the other hand, the k-NN classification process has 83.3% of accuracy, since 250 instances have been correctly classified .

5 Conclusion

This paper presents a fuzzy system construction to the automatic Latin musical genres classification. For each one of the ten musical genres considered (tango, salsa, forró, axé, bachata, bolero, merengue, gaúcha, sertanejo and pagode), a fuzzy classification system is constructed for each one of the three input groups, which are: beat related (which includes the relative amplitudes and the beats per minute), timbre texture (the first five MFCCs) and pitch related (which includes the maximum periods and amplitudes of the pitch histograms). Then, a membership percentage of an instance is obtained using the arithmetic mean of the three outputs.

A given instance, which contains the numeric inputs described in Table 1, can be inserted into the fuzzy classification systems of each one of the ten musical genres. As a result, the membership percentage of this given instance in each musical genre is obtained with 90.7% of success (accuracy). Moreover, this membership percentage demonstrates that the instance can be a fusion of two or more musical genres. The k-NN classification, one of the most traditional classification algorithm, presents 83.3% of accuracy. It is relevant to say that this method has become obsolete, since the traditional classifiers do not represent the musical genres fusion. By comparing with the human classification, not only is it not an automatic process, but it also has uncertainties about the decision making. The increasing number of musical genres and their fusion motivate the use of fuzzy logic, due to the uncertainties and the fuzziness among genres boundaries.

References

[1] A. Berenzweig, B. Logan, D. P. Ellis, and B. Whitman. A largescale evaluation of acoustic and subjective music-similarity measures. *Computer Music Journal*, 28:63–76, 2004.

[2] D. Correa and F. Rodrigues. A survey on symbolic data-based music genre classification. *Expert Systems with Applications*, 60.

[3] Y. M. G. Costa, L. E. S. Oliveira, and C. N. Silla Jr. An evaluation of convolutional neural networks for music classification using spectrograms. *Applied Soft Computing*, 52:28–38, 2017.

[4] R. Dannenberg, J. Foote, G. Tzanetakis, and C. Weare. Panel: new directions in music information retrieval. *Proc. International Computer Music Conference, Havana, Cuba*, 2001.

[5] R.O Duda, P. E Hart, and D. G. Stork. *Pattern Classification.* Wiley, 2001.

[6] F. Fernández and F. Chávez. On the application of fuzzy rule-based systems to musical genre classification. *1st Workshop on Evolutionary Music. IEEE CEC, New Orleans, EE.UU*, pages 25–31, 2011.

[7] F. Fernández and F. Chávez. Fuzzy rule based system ensemble for music genre classification. *Evolutionary and Biologically Inspired Music, Sound, Art and Design. Lecture Notes in Computer Science*, 7247:84–95, 2012.

[8] F. Fernández, F. Chávez, R. Alcalá, and F. Herrera. Musical genre classification by means of fuzzy rule-based systems: A preliminary approach. *IEEE Congress on Evolutionary Computation, IEEE CEC, New Orleans, EE.UU*, pages 2571–2577, 2011.

[9] Z. Fu, G. Lu, K.M. Ting, and D. Zhang. A survey of audio-based music classification and annotation. *IEEE Transactions on Multimedia*, 13(2):303–319, 2011.

[10] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka. Rwc music database: Popular, classical and jazz music databases. *Proc. of the International Conference on Music Information Retrieval*, 2002.

[11] J. R. Jang. Anfis: Adaptive - network - based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3):665–685, 1993.

[12] T. Lidy and A. Rauber. Evaluation of feature extractors and psycho-acoustic transformations for music genre classification. *Proc. of the Sixth International Conference on Music Information Retrieval*, 2005.

[13] G. Zoia N. Scaringella and D. Mlynek. Automatic genre classification of music content: A survey. *IEEE Signal Processing Magazine*, 23(2):133–141, 2006.

[14] T. Nakashika, C. Garcia, and T. Takiguchi. Local-feature-map integration usingconvolutional neural networks for music genre classification. *Interspeech*, 2012.

[15] W. E Pedrycz and F. Gomide. *An Introduction to Fuzzy Sets.* MIT Press, 1998.

[16] P. Ponce León, Rizo D., R. Ramírez, and J. M. Iñesta. Melody characterization by a genetic fuzzy system. *5th Sound and Music Computing Conference, Berlin, Germany*, pages 15–20, 2008.

[17] D. Rizo, P. J. Ponce De León, C. Pérez-Sancho, A. Pertusa, and J. M. Iñesta. A pattern recognition approach for melody track selection in midi files. *Proc. 7th International Symposium on Music Information Retrieval ISMIR, T. A. Dannenberg R., Lemstrom K., Ed., Victoria, Canada*, pages 61–66, 2006.

[18] S. Sigtia and S. Dixon. Improved music feature learning with deep neuralnetworks. *Proc. of the IEEE International Conference on Acoustic, Speech and SignalProcessing*, 2014.

[19] C.N. Silla Jr, A.L. Koerich, and C.A.A Kaestner. The latin music database. *Proc. International Society for Music Information Retrieval*, pages 451–456, 2008.

[20] C.N. Silla Jr, A.L. Koerich, and C.A.A Kaestner. A feature selection approach for automatic music genre classification. *International Journal of Semantic Computing*, 3(2):183–208, 2009.

[21] B. Sturm. An analysis of the gtzan music genre dataset. *Proc. of the 2nd Int. ACM Workshop on Music Information retrieval with user-centered and multimodal strategies.*

[22] G. Tzanetakis and P. Cook. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5):293–302, 2002.

[23] G Tzanetakis and Perry Cook. Marsyas: A framework for audio analysis. *Organized Sound*, 4:169–175, 1999.

[24] G Tzanetakis and Perry Cook. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5):293–302, 2002.

[25] L. A. Zadeh, K. S. Fu, K. Tanaka, and M. Shimura. *Fuzzy sets and their applications to cognitive and decision processes.* Academic Press, 1975.

[26] H. J. Zimmermann. *Fuzzy sets theory and its applications.* Kluwer, 1991.