ISSN:0975-9646

Classification of Music based on Clave Direction using Machine Learning Techniques

Sabina M A

Department of Computer Science and Engineering, TKM College of Engineering, Kollam,Kerala, India

***Abstract*— Application of computational intelligence in music is enormous in the modern era. Clave direction, a rhythmic–regulative principle in Afro Latin music, is difficult for music experts to explain why clave works the way it does. Accurate use of clave is a complex mathematical and psycho acoustical task that one can learn only after extensive exposure and practice assisted by technological clave-training devices. Clave-aware training and recording technology can have far reaching effects in audio recording, notation and sequencing, auto-accompaniment keyboards and music query systems. In this paper, the problem of classification of music converted into 16 bit binary attack-point vectors into their corresponding clave-direction classes is addressed using machine learning techniques such as linear discriminant analysis and feed forward neural networks using back propagation algorithm.**

Keywords— **clave direction, lda, backpropagation**

I. INTRODUCTION

Today the application of computational intelligence in music is enormous. Automatic classification of music, style-based interpreter recognition, automatic composition and improvisation, music recommender systems, genre and tag prediction ,score alignment, polyphonic pitch detection, chord extraction, pattern discovery, beat tracking, expressive performance modelling are some of the domains where computation intelligence plays a major role

Clave direction is a concept and a rhythmic–regulative principle. It is analogous to the key of a piece of music, but instead of governing tonality, it governs fine–scale local timing.[1] [2]Clave direction gives a composer or improviser a set of preferred timing options and the appropriate places within each phrase to place them (a sort of micro–phrasing), along with acceptable ways to break these rules. Just as the key determines the tonal center of the piece, the clave direction is the overarching determinant of the timing preferences for the piece, but like its tonal counterpart, it allows for variety in musical expression, and even tricks, puzzles and multi–layered playfulness in its execution.[1][2]

The expert musicians consider it important for musicians involved in Afro-Latin music to be able to abide with the unwritten rules of the music with the expertise close to that of practitioners who have been brought up immersed in these cultural subtleties. 1][2][3] Thus a thorough understanding of these rules is necessary for anyone who is involved in Afro Latin music[1][2][3][4]

The traditional mode of learning involves years of exposure beginning very early in life like that of classical music in India. However, psychological study on rhythm suggests that rhythmic awareness arises after the first year of life. Either way, early exposure appears to be crucial in clave enculturation.[1]So those who wish to learn this kind of music including outsiders of this culture find it difficult to grasp its intricacies.[1] It is difficult to explain why clave works the way it does for musicians who have clave-based music as part of their culture.[1][2].It is often described as knowledge that is passed through generations.[1]. So it would be of great help to those who doesn't have long time for learning and for outsiders to the culture if electronic devices that can impart expert knowledge to end-users through state-of-the-art digital signal processing (DSP) and CI techniques are developed. Similarly, the recognition and idiomatically accurate use of clave is a complex mathematical and psycho acoustical task that one can learn only after extensive exposure and practice – a process that can be accelerated and assisted by technological clave-training devices[1].Once developed, clave-aware training and recording technology could be put to use in audio recording, notation and sequencing, auto-accompaniment keyboards, music query systems and various types of electronic music coaches, existing or to be developed.[1][2][3]

II. RELATED WORK

In the work by Mehmet Vurkac et al, prestructuring multilayer perceptron based on Information-Theoretic Modeling of a Partido-Alto-based Grammar for Afro-Brazilian Music has been done through domain knowledge built into the solution through reconstructability analysis, a form of information-theoretic modeling, which is used to identify mathematical models that can be transformed into a graphic representation of the problem domain's underlying structure.[1].Prestructuring reduces the set of all possible map that are realizable by the neural network.[1]

In[3] Mehmet Vurkac presents an in depth study of the *clave direction* The main points made in this article are that

1) the absolute onset count should be abandoned;

2) upbeats and downbeats are on beats as far as clave direction is concerned;

3) the inner/outer demarcation should be preferred over the standard approach of dividing phrases into a first and second half; and it is the relative off beatness of consecutive sections that determines the sense of clave, not counting onsets or template–matching against well–known patterns.[1][2][3]

III. PROBLEM

In a typical Afro Latin music-making scenario, musicians specify the overall clave direction for a piece being composed, recorded or sequenced just as ones specifies a tempo, key and meter [1]. If a clave-analysis tool can be

developed, it could then flag occurrences in the recorded, notated or sequenced music that go counter to the chosen clave direction (cross clave). [1].Setting'vigilance' or 'strictness' parameter or a cultural-context variable is another option [1].
The classification of binary attack point vectors according to their clave directions is addressed here using two machine learning techniques namely Linear discriminant Analysis (LDA) and Neural Network.

A. LINEAR DISCRIMINANT ANALYSIS(LDA) :

LDA seeks to reduce dimensionality while preserving as much of the class discriminatory information as possible. Fishers criterion tries to find the projection that maximises the variance of the class means and minimises the variance of the individual classes Fisher's linear discriminant is given by the vector 'w' which maximizes

$$J(w) = W^T S_B W / W^T S_W W$$

where S_B is the between class scatter matrix and S_W is the within class scatter matrixThe steps in LDA are as follows

1. Compute the *d*-dimensional mean vectors for the different classes from the dataset.
2. Compute the scatter matrices (between-class and within-class scatter matrix).
3. Compute the eigenvectors ($\mathbf{e_1}, \mathbf{e_2}, ..., \mathbf{e_d}$) and corresponding eigenvalues ($\lambda_1, \lambda_2, ..., \lambda_d$) for the scatter matrices.
4. Sort the eigenvectors by decreasing eigenvalues and choose **k** eigenvectors with the largest eigenvalues to form a $d \times k$-dimensional matrix **W** (where every column represents an eigenvector).
5. Use this $d \times k$ eigenvector matrix to transform the samples onto the new subspace. This can be summarized by the equation $\boldsymbol{Y} = \boldsymbol{X} \times \boldsymbol{W}$ (where $\boldsymbol{X}$ is an $n \times d$-dimensional matrix; the ith row represents the i^{th} sample, and $\boldsymbol{Y}$ is the transformed $n \times k$-dimensional matrix with the n samples projected into the new subspace).[7]

B. FEED-FORWARD MULTILAYER NEURAL NETWORK

The back propagation algorithm trains a given feed-forward multilayer neural network for a given set of input patterns with known classifications. When each entry of the sample set is presented to the network, the network examines its output response to the sample input pattern. The output response is then compared to the known and desired output and the error value is calculated. Based on the error, the connection weights are adjusted. The back propagation algorithm is based on Widrow-Hoff delta learning rule in which the weight adjustment is done through mean square error of the output response to the sample input . The set of these sample patterns are repeatedly presented to the network until the error value is minimized.[6][8]

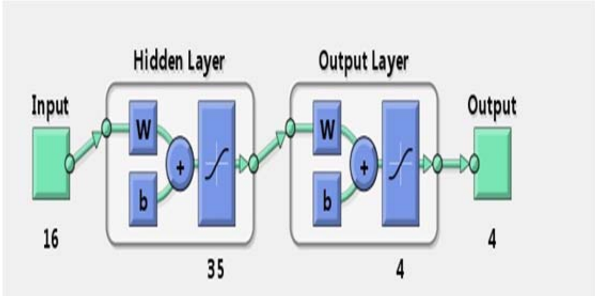

Fig 1. Layout of the Neural Network used

There are two phases in back propagation algorithm.
Phase 1: Propagation

1. Forward propagation of a training pattern's input through the neural network in order to generate the propagation's output activations.
2. Backward propagation of the propagation's output activations through the neural network using the training pattern target in order to generate the deltas of all output and hidden neurons
.

Phase 2: Weight update
For each weight-synapse follow the following steps:

1. Multiply its output delta and input activation to get the gradient of the weight.
2. Subtract a ratio (percentage) of the gradient fr the weight.

This ratio (percentage) influences the speed and quality of learning; it is called the learning rate. The greater the ratio, the faster the neuron trains; the lower the ratio, the more accurate the training is. The sign of the gradient of a weight indicates where the error is increasing; this is why the weight must be updated in the opposite direction. Repeat phase 1 and 2 until the performance of the network is satisfactory.[6][[8]

IV. EXPERIMENTS

The clave direction classification data set has been downloaded from UCI machine learning repository. [6][3][4][5] The data samples are binary attack-point vectors and their clave-direction classes according to the partido-alt1-based paradigm.[1][3][4][5].
The data consist of 16 binary inputs and one 'four-bit' one-hot classification output. The 16-bit inputs are binary-valued attack-point vectors. 1 indicates the substantial presence (0, absence) of an onset (note start) in a certain time window during one bar of 4/4 time music (not limited to percussion, hence onset vectors without duration) quantized to 16th-note subdivisions. Each vector has 16 positions in which there may be or not be an onset. The output classes (left to right: neutral, reverse clave, forward clave, and incoherent) were determined through the music-theoretic/ethno musicological portion of the studies, based on both double-blind listening tests and informal interviews with four professional master-musicians, as well as decades of studying the music[2][3][4][5].
In terms of divisive rhythm counting, the first 16 attributes (input bits) correspond to a significant onset of one bar of 4/4 time. The last four are the output classes

(3 - neutral, 2 - reverse clave, 1 - forward clave, 0 - incoherent) in one-hot (one-up) encoding. [2][3][4][5]

Experiments have been conducted over samples in the dataset with both lda and neural network

A multi layer perceptron feed forward network with back propagation has been used for training, validation and testing. Using 35 hidden layer neurons, training was successfully done .For training 70% of random samples were used and 15% each were used for validation and testing

Using multiclass linear discriminant analysis, the classification has been carried out with the same set of input used for neural network and the performance has been studied

V. Conclusions

The classification of data samples which are binary attack-point vectors and their clave-direction class according to the partido-alto-based paradigm has been keenly studied and experimented using linear discriminant analysis. Experimentations showed that Multi layer perceptron neural network outperforms linear discriminant analysis when the misclassification rate is studied This classification study has to be extended using other machine learning classification techniques like svm .Hence these experimentations will greatly help in the development of a clave aware music system as this genre is highly popular in Latin America and Africa and is very difficult to learn. The partido-alto-based paradigm has been keenly studied and experimented using linear discriminant analysis .Experimentations showed that Multi layer perceptron neural network outperforms linear discriminant analysis when the misclassification rate is studied

This classification study has to be extended using other machine learning classification techniques like svm .Hence these experimentations will greatly help in the development of a clave aware music system as this genre is highly popular in Latin America and Africa and is very difficult to learn.

Table I
Sample Of Dataset Used

INPUT (binary attack-point vectors)	CLASS	TYPE (clave direction)
1 1 1 1 1 1 1 1 0 0 1 0 1 1 1 0	0 0 0 1	incoherent
1 1 1 1 1 1 1 0 0 1 1 0 0 1 0 0	0 0 0 1	incoherent
0 1 0 1 0 0 0 0 1 0 0 1 0 0 0 1	0 0 1 0	forward clave
0 1 0 1 0 0 0 0 0 1 1 0 0 0 0 1	0 0 1 0	forward clave
1 0 0 1 1 1 0 1 0 1 1 1 0 0 1 1	0 1 0 0	reverse clave
1 0 0 1 1 1 0 1 0 1 1 1 0 0 1 0	0 1 0 0	reverse clave
1 1 1 0 0 0 1 1 1 1 1 1 1 0 0 1	1 0 0 0	neutral
1 1 0 0 1 1 0 0 1 0 0 0 0 0 0 0	1 0 0 0	neutral

Table Ii
Classification Accuracy

Method	Classification Accuracy%
Feed forward Neural Network using back propagation algorithm	98.3
LDA Analysis	90.5

Acknowledgement

I sincerely express my heartfelt gratitude to Almighty and to everyone who supported me to prepare this paper. I am thankful for the aspiring guidance, invaluably constructive criticism and advice from everyone during this work. I am sincerely thankful to all who shared their opinions on all issues related to the work.

References

[1] Vurkaç, Mehmet, "Prestructuring Multilayer Perceptrons based on Information-Theoretic Modeling of a Partido-Alto-based Grammar for Afro-Brazilian Music: Enhanced Generalization and Principles of Parsimony, including an Investigation of Statistical Paradigms," *Dissertations and Theses.* Paper 384,2011

[2] Vurkaç, Mehmet, "On the Need for Clave-Direction Analysis: A New Arena for Educational and Creative Applications of Music Technology ," Journal of Music, Technology and Education, Vol. 4, No.1,August2011

[3] Vurkaç, Mehmet, ,"A Cross-Cultural Grammar for Temporal Harmony in Afro-Latin Musics: Clave, Partido-Alto and Other Timelines Current Musicology," No. 94, 2014 with Fall 2012

[4] *Mehmet Vurkaç*, "*Balanço*: The Contour of Relative Offbeatness,'" Proceedings of Bridges 2012: Mathematics, Music, Art, Architecture,CulturePages,329–336 (2012)

[5] (2015)The UCI machine learning repository [Online] Available :http://archive.ics.uci.edu/ml

[6] Rumelhart, David E., Hinton, Geoffrey E.; Williams, Ronald J,"Learning representations by back-propagating errors". *Nature* **323** (6088): 533–536.,8 October 1986.

[7] (2014)LDA [Online] Available :http:// spartanideas.msu.edu

[8] (2015) wiki []http://en.wikipedia.org/