

A COMPARISON OF RHYTHMIC SIMILARITY MEASURES

Godfried T. Toussaint
McGill University
School of Computer Science

ABSTRACT

Traditionally, rhythmic similarity measures are compared according to how well rhythms may be recognized with them, how efficiently they can be retrieved from a data base, or how well they model human perception and cognition. In contrast, here similarity measures are compared on the basis of how much insight they provide about the structural inter-relationships that exist within families of rhythms, when phylogenetic trees and graphs are computed from the distance matrices determined by these similarity measures. Phylogenetic analyses yield insight into the evolution of rhythms and may uncover interesting *ancestral* rhythms.

1. INTRODUCTION

Measuring the similarity between rhythms is a fundamental problem in computational music theory, with many applications such as music information retrieval and copyright infringement resolution. It has been argued that the type of similarity measure used should depend on the intended application [8]. We describe the results of a pilot study carried out to compare a variety of rhythm similarity measures with respect to how useful they are for the phylogenetic analysis of families of rhythms. It should be emphasized that the focus of this paper is not the measurement of psychological similarity (based on perception). It is also assumed that the data is obtained from musical scores rather than audio, and thus tempo problems are ignored. Two families of rhythms were selected for this study, the fundamental rhythm timelines in 4/4 and 12/8 time, used in traditional West-African and Afro-American music [12], [14]. These timelines are also called bell-patterns or *clave* patterns, of which the most famous on the world music scene is the *clave Son*. Although the data used in this study consist of only two small families of rhythms, these were chosen for two reasons: (1) they are important and well known in world music, and (2) they are simple and well understood from the musicological point of view, thus simplifying the validation process of their phylogenetic analyses.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.
© 2004 Universitat Pompeu Fabra.

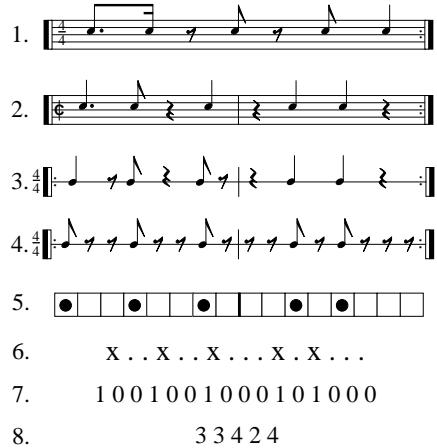


Figure 1. Eight common ways of representing the *clave Son* rhythm.

The *clave Son* is usually notated using standard music notation which affords many ways of expressing a rhythm. Four such examples are given in the top four lines of Figure 1. Western music notation is not ideally suited to represent African music [1]. The remaining lines show ways of representing rhythms that avoid Western notation.

There exists a wide variety of methods for measuring the similarity of two rhythms represented by strings of symbols [13]. Traditionally similarity between two patterns is measured by a simple template matching operation. More recently similarity has been measured with more powerful and complex functions such as the *earth mover's distance* [2].

In this section we describe the five, easy to compute, popular and representative measures of rhythm similarity (or dissimilarity) that were compared in this study. It should be noted that although the rhythms considered here have, within each family, the same number of onsets, the Hamming, swap, and chronotonic distance measures do not assume this restriction and work just as well for comparing rhythms containing unequal numbers of onsets.

The Hamming distance: The Hamming distance is the number of places in the binary n -bit string representation of the rhythm where bits do not match. It is easily computed in $O(n)$ time.

The Euclidean interval vector distance: Some rhythm detection algorithms and systems for machine recognition of music patterns [3] use *inter-onset intervals* as the basic

features for measuring similarity. These are the intervals of time between consecutive note onsets in a rhythm. The dissimilarity between two rhythms, each consisting of n time intervals, may be computed in $O(n)$ time using the Euclidean distance between the two interval vectors.

The interval-ratio distance: Coyle and Shmulevich [3] represent a music pattern by what they call a *difference-of-rhythm vector*. If $T = (t_1, t_2, \dots, t_n)$ is a vector of inter-onset time intervals then they define the difference-of-rhythm vector as $X = (x_1, x_2, \dots, x_{n-1})$, where $x_i = t_{i+1}/t_i$. They propose a distance measure based on this vector which may also be easily computed in $O(n)$ time. This measure is called here the *interval-ratio distance*.

The swap distance: A generalization of the Hamming distance is the *edit* distance which allows for insertions and deletions of notes [9]. Using *dynamic programming* the edit distance may be computed in $O(n^2)$ time.

The problem of comparing two binary strings of the same length with the same number of one's suggests an extremely simple edit operation called a *swap*. A swap is an interchange of a one and a zero that are adjacent to each other. The swap distance is the *minimum* number of swaps required to convert one rhythm to another. Note that this distance measure extends naturally to handle rhythms with a different number of onsets. Assume that rhythm A has more onsets than rhythm B . Then the swap distance may be defined as the minimum number of swaps required to take *all* onsets of A to the onsets of B , with the restriction that each onset of B must receive at least one onset of A , as was done in [4] for the analysis of Flamenco rhythms.

The swap distance may be computed by actually performing the swaps, but this is inefficient because a quadratic number of swaps may be required in the worst case. If we compare the interval distances instead, the swap distance may be computed in $O(n)$ time when the rhythms have the same number of onsets, and in some other situations as well, resulting in a large gain over using linear or dynamic programming.

The chronotonic distance: In 1987 Kjell Gustafson, at the Phonetics Laboratory of the University of Oxford, proposed an original method to represent a rhythm as a two-dimensional graph [5]. His idea is best explained with an example. Consider first the clave *Son* pattern shown on line five of Figure 1. Although this representation is realistic in terms of the time at which beats occur, the relative durations of the intervals are not easily observed. In a *histogram* approach to rhythm visualization the intervals are plotted along the y -axis [5], resulting in the *adjacent-interval-spectrum* of the rhythm in which the relative lengths of the intervals are clearly visible but the temporal information along the x -axis is lost. To obtain a graphical representation that possesses the advantages of both of these methods, Gustafson simply combines them. The result of this union is illustrated with the clave *Son* in Figure 2. Each temporal element (interval) is now a box and both the x and y axes represent the length of time of the interval. Gustafson refers to such a display as TEDAS (*Temporal Elements Displayed As Squares*). This

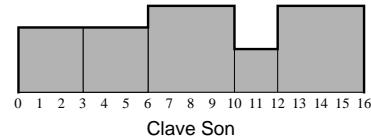


Figure 2. The *chronotonic* (TEDAS) representation of the clave *Son*.

idea was rediscovered by Hofmann-Engl [6]. We will refer to this as the *chronotonic* representation of a rhythm.

Viewing the chronotonic representation as a piece-wise linear function opens the door to a large family of possible distance functions with a long history in the fields of statistics and pattern recognition [10]. Given two probability density functions $f_1(x)$ and $f_2(x)$, there are many measures of the distance (or separation) between them. One of the most well known is the *Kolmogorov variational* distance [11] given by:

$$K = \int |f_1(x) - f_2(x)| dx. \quad (1)$$

Here the measure K is used to compare rhythms using the chronotonic representation proposed by Gustafson [5] (even though the functions are not probability density functions). In this discrete setting it is clear that K may be computed easily in $O(n)$ time.

2. COMPARING RHYTHMIC SIMILARITY MEASURES

The measures discussed in the preceding section were compared using two families of rhythms: the fundamental binary and ternary timelines used in West-African and Afro-American traditional music. The first family consists of the six five-beat clave-bell patterns in 4/4 time [12], and the second is comprised of the ten seven-beat bell patterns in 12/8 time [14]. These two families are depicted in the chronotonic notation of Gustafson in Figures 3 and 4, respectively. One of the main goals of this study is to obtain insight into the evolution of rhythms as well as uncover interesting “ancestral” rhythms. Hence, rather than use any of the myriad traditional cluster analysis methods, here phylogenetic analysis tools from bioinformatics are used. After all, at a purely mathematical level, rhythms and DNA molecules are both sequences of symbols.

Several techniques exist for generating phylogenetic trees from distance matrices [7]. Some of these methods have the desirable property that they produce graphs that are *not* trees, when the underlying structure is inherently *not* tree-like. One notable example is *SplitsTree* [7]. Like the more traditional phylogenetic trees, *SplitsTree* computes a plane graph embedding with the property that the distance in the drawing between any two nodes reflects the true distance between the corresponding rhythms in the distance matrix. However, if the tree structure does not

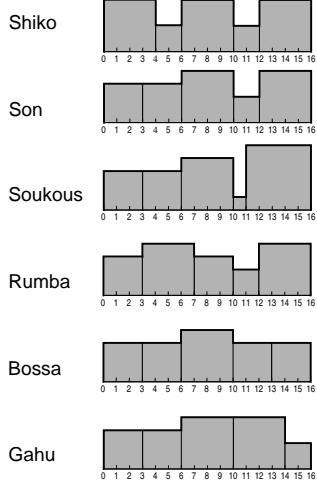


Figure 3. The chronotonic representation of the 4/4 time *clave* patterns.

match the data perfectly then new nodes in the graph may be introduced with the goal of obtaining a better fit. Such nodes may suggest implied “ancestral” rhythms from which their “offspring” may be derived. In addition, edges may be split to form parallelograms, such as in Figure 5. The relative sizes of these parallelograms are proportional to an *isolation index* that indicates how significant the clustering relationships inherent in the distance matrix are. *SplitsTree* also computes the *splittability index*, a measure of the goodness-of-fit of the entire splits graph. This *fit* is obtained by dividing the sum of all the approximate distances in the splits graph by the sum of all the original distances in the distance matrix [7]. For these reasons *SplitsTree* was the method of choice in this study.

3. RESULTS

Space limitations do not permit a discussion of all the results. Suffice it to say that the quantitative phylogenetic analyses support the tenets previously established for these rhythms via traditional qualitative musicological methods. Here only some of the results obtained with the best measure, the chronotonic distance, are summarised.

For the 4/4 time rhythms the splits graph shown in Figure 5 is a tree-like graph with a perfect fit of 100%. This distance measure is the only one that exhibits a strong clustering in this family of rhythms. In particular, the long parallelogram clearly separates *Bossa-Nova* and *Gahu* from the rest. An examination of their chronotonic functions of time (see Figure 3) yields a clearly distinguishable discriminating feature. The *Bossa-Nova* and *Gahu* functions have only *one* local maximum, whereas each of the four other rhythms has *two* local maxima.

For the 12/8 time rhythms the splits graph shown in Figure 6 exhibits a rich structure with a perfect fit of 100%. There are two highly clustered groups. The first group

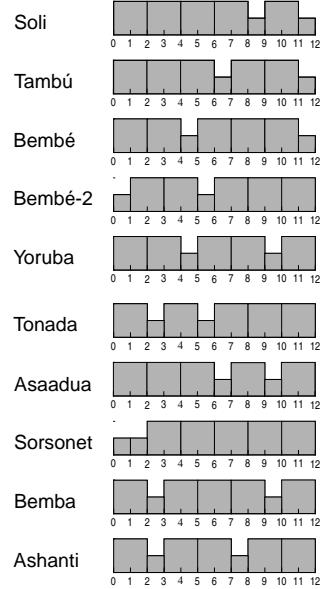


Figure 4. The chronotonic representation of the 12/8 time *bell* patterns.

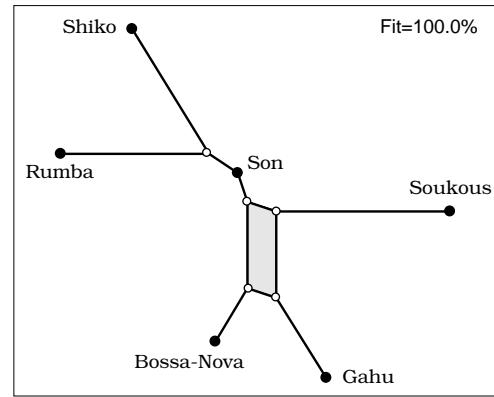


Figure 5. The 4/4 time rhythms.

consists of *Soli*, *Asaadua*, *Tambú*, *Yoruba*, and *Bembé*. The second group is comprised of *Tonada*, *Ashanti*, *Bembé-2*, and *Sorsonet*. Furthermore, *Bemba* lies half-way between the two groups.

The splits graph has nine potential ancestral nodes. One of these is distinguished from the rest, not only because of its high degree, but because it is located in the center of the entire collection, and is at a distance of *two* from *all* ten rhythms. For this reason it is labelled the *2-center* of the set. A reconstruction of this ancestral rhythm yields the pattern $[x \cdot x \cdot x \cdot x \cdot x \cdot x \cdot]$, which is a steady, isochronous “heart-beat.”

Several desirable criteria may be used for comparing phylogenetic trees of families of rhythms. Such criteria include (1) simplicity, (2) goodness of fit, (3) indication of clustering, and (4) generation of interesting “ancestral” rhythms.

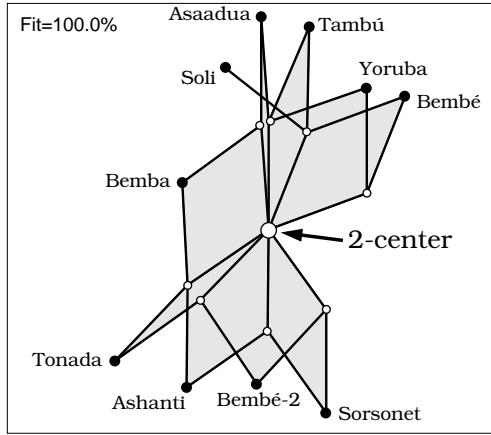


Figure 6. The 12/8 time rhythms.

Simplicity: The simplest splits graphs were obtained with the swap distance.

Goodness of fit: The only distance measures that yielded a 100% fit for both binary and ternary rhythms were the swap and chronotonic distances. The Hamming distance also produced a 100% fit for the binary rhythms.

Clustering: The most impressive clustering was obtained with the chronotonic distance. For the binary rhythms, *Gahu* and *Bossa-Nova* are clearly differentiated from the rest. For the ternary rhythms the splits graph yields two well separated clusters.

Ancestral rhythm generation: The most noteworthy “ancestral” rhythms were produced with the chronotonic distance for ternary rhythms and with the Hamming distance for the binary rhythms.

In conclusion, a comparison of the significance of the role played by the five distance measures according to each of the four criteria outlined in the preceding, suggests that the best rhythmic dissimilarity measure is the *chronotonic* distance, followed by the *swap* distance in close second place. Generalizing these conclusions to other larger families of rhythms is premature. However, these results constitute a positive first step, and encourages the continuation of this line of research in greater depth.

4. REFERENCES

- [1] Simha Arom. *African Polyphony and Polyrhythm*. Cambridge University Press, Cambridge, England, 1991.
- [2] Sung-Hyuk Cha and Sargur N. Srihari. On measuring the distance between histograms. *Pattern Recognition*, 35:1355–1370, 2002.
- [3] E. J. Coyle and I. Shmulevich. A system for machine recognition of music patterns. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Seattle, Washington, 1998.
- [4] Miguel Díaz-Bañez, Giovanna Farigu, Francisco Gómez, David Rappaport, and Godfried T. Toussaint. El compás flamenco: a phylogenetic analysis. In *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science*, Southwestern College, Winfield, Kansas, July 30 - August 1 2004.
- [5] Kjell Gustafson. The graphical representation of rhythm. In *(PROPH) Progress Reports from Oxford Phonetics*, volume 3, pages 6–26, University of Oxford, 1988.
- [6] Ludger Hofmann-Engl. Rhythmic similarity: A theoretical and empirical approach. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, and J. Renwick, editors, *Proceedings of the Seventh International Conference on Music Perception and Cognition*, pages 564–567, Sidney, Australia, 2002.
- [7] Daniel H. Huson. SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics*, 14:68–73, 1998.
- [8] Benoit Meudic. Musical similarity in a polyphonic context: a model outside time. In *Proceedings of the XIV Colloquium on Musical Informatics*, Firenze, Italy, May 8-10 2003.
- [9] Keith S. Orpen and David Huron. Measurement of similarity in music: A quantitative approach for non-parametric representations. In *Computers in Music Research*, volume 4, pages 1–44. 1992.
- [10] Godfried T. Toussaint. On the divergence between two distributions and the probability of misclassification of several decision rules. In *Proceedings of the Second International Joint Conference on Pattern Recognition*, pages 27–35, Copenhagen, Denmark, August 13-15 1974.
- [11] Godfried T. Toussaint. Sharper lower bounds for discrimination information in terms of variation. *IEEE Transactions on Information Theory*, pages 99–100, January 1975.
- [12] Godfried T. Toussaint. A mathematical analysis of African, Brazilian, and Cuban *clave* rhythms. In *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science*, pages 157–168, Towson University, Towson, MD, July 27-29 2002.
- [13] Godfried T. Toussaint. Algorithmic, geometric, and combinatorial problems in computational music theory. In *Proceedings of X Encuentros de Geometría Computacional*, pages 101–107, University of Sevilla, Sevilla, Spain, June 16-17 2003.
- [14] Godfried T. Toussaint. Classification and phylogenetic analysis of African ternary rhythm timelines. In *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science*, pages 25–36, Granada, Spain, July 23-27 2003.