ON THE FEASIBILITY OF MARKOV MODEL BASED ANALYSIS OF GEORGIAN VOCAL POLYPHONIC MUSIC

Frank Scherbaum
UP Transfer,
University of Potsdam
fs@geo.uni-potsdam.de

Simha Arom
UMR 7206, CNRS-MNHN, Paris,
simha.arom@gmail.com

Frank Kane
Kane.frank@gmail.com

1. INTRODUCTION

Here we present the first results of a feasibility study regarding the usefulness of Markov models (e. g. Meyn and Tweedie, 2005) for the analysis of Georgian vocal polyphonic music. The present analysis, which can be seen as an attempt to explore new tools for the syntactic analysis of vertical concomitances in music (c.f. Arom & Vallejo, 2008; 2010), is based on the collection of folk songs from Svaneti (NW Georgia) described by Akhobadze (1957). For this purpose, we developed a software package in Mathematica 10 (Wolfram Research, 2014), dedicated to musical score analysis.

2. METHODOLOGICAL FRAMEWORK

In the context of the present study we make the assumption that a song can approximately be treated as a temporal process in which discrete harmonic or melodic states change according to probabilistic rules which are implicitly contained in the song itself. In more technical terms, the assumption made is that the sequence of "chords" (or more precisely vertical conjunctions) in a song can be modeled as a discrete Markov chain of harmonic states (e. g. multi-voice chords of particular duration). In this context, the probability that a particular state changes into another one is assumed to depend only on the current state and a state transition probability matrix T (Markov property), which can be estimated from the musical score by statistical analysis.

3. PROCESSING

In order to determine the Markov model for a Georgian song, the score (cf. Fig. 1) is first converted into digital form (musicXML). Subsequently, the pitches in each voice (in Georgian polyphony usually three) are approximated by piecewise functions of time (Fig. 2), which subsequently are combined into a vector function. This enables the easy algorithmic determination of the margins of harmonic states as times where the time-derivatives of this function are non-zero.

Figure 1. Original score of the song *Lushnu Läshgaru* (Svan marching song).

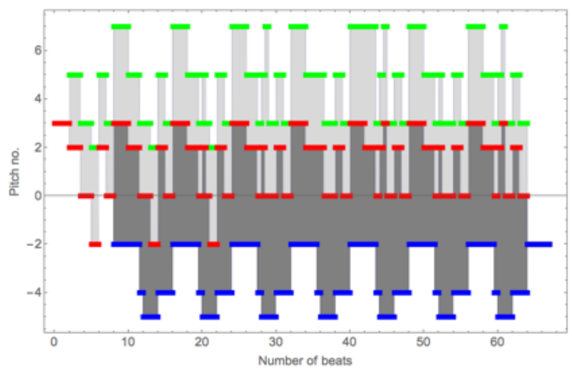

Figure 2. Pitch function of the song *Lushnu Läshgaru*. The red, green and blue lines correspond to the pitches (in pitch number) for the middle, top and bass voice of the song, respectively. These can be seen as components of a vector function, the non-zero derivatives of which define boundaries of the harmonic states of the song.

Since Georgian music is considered modal, we determined the mode of each song (see Figs. 3 and 4) as the mode which is most consistent with the pitch inventory of the song, and converted all chords into mode-degree representations. Regarding the reference note, we followed the strategy of Arom & Vallejo (2008) to use the *finalis*.

Figure 3. Example for the determination of the song´s mode based on the statistical analysis of the pitch inventory of *Lushnu Läshgaru*. The histograms show the distribution of fractional (regarding the duration of the song) consistency (to the right) and inconsistencies (to the left) of the pitch inventories of the individual voices. The center panel shows that the complete pitch inventory of *Lushnu Läshgaru* for all voices is fully consistent with the mode Sol (G), but none other.

Total	Top	Middle	Bass	Mode
0	0	0	0	Sol
0.291667	0	0	0.291667	Do
0.375	0.0333333	0.341667	0	Re
0.841667	0.3	0.25	0.291667	Fa
0.891667	0.316667	0.341667	0.233333	La
1.2	0.316667	0.65	0.233333	Mi
1.51667	0.633333	0.65	0.233333	Si

Figure 4. Mode inconsistency table of the pitch inventory of *Lushnu Läshgaru*. The entries in the different columns indicate the fractional inconsistencies in all voices combined (column 1) and the individual voices (columns 2 – 4) for the tested modes (indicated in column 5). As can be seen from the total inconsistency value of 0, that the complete pitch inventory of *Lushnu Läshgaru* for all voices is fully consistent with the mode Sol (G). For the mode Do (C), there would be inconsistencies with the pitches in the bass melody.

Subsequently, the state transition matrix T (Fig. 5) is determined by statistical analysis of the actually occurring state transitions in the song.

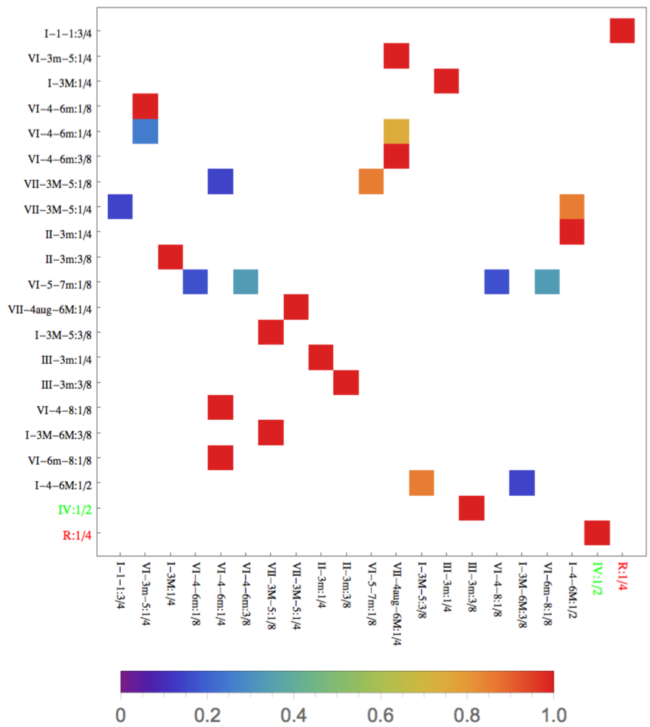

Figure 5. State transition matrix T of the complete song *Lushnu Läshgaru*.

The color code of an element of T describes the relative frequency (loosely speaking the probability) with which the state indicated by the row label is changing into the state indicated by the column label. Elements which are white, correspond to state transitions with zero probability. For example the 3rd row from the bottom corresponds to the state I-4-6M:1/2. It contains two non-zero elements, on for the transition to the state I-3M-5:3/8 with transition probability close to 0.8 (orange color) and one for the transition to the state I-3M-6:3/8 with transition probability close to 0.2 (blue color). Elements with red color describe transitions which take place with probability of 1. In other words, if song comes to the state described by the row label for a red element, the next state is fully determined.

From T we generated the so-called process graphs (Fig. 6) which within the Markov model framework contains the complete information about the corresponding song in a visual (but still quantitative) way. Each node corresponds to a harmonic state, while the edges of the graph represent possible pathways through the harmonic state inventory with their corresponding probabilities color coded as shown in the legend. The start and end states of the song are indicated by green and red text colors, respectively. Process graphs are convenient tools to visually compare songs with respect to their complexity and hence to the predictability of their chord progressions.

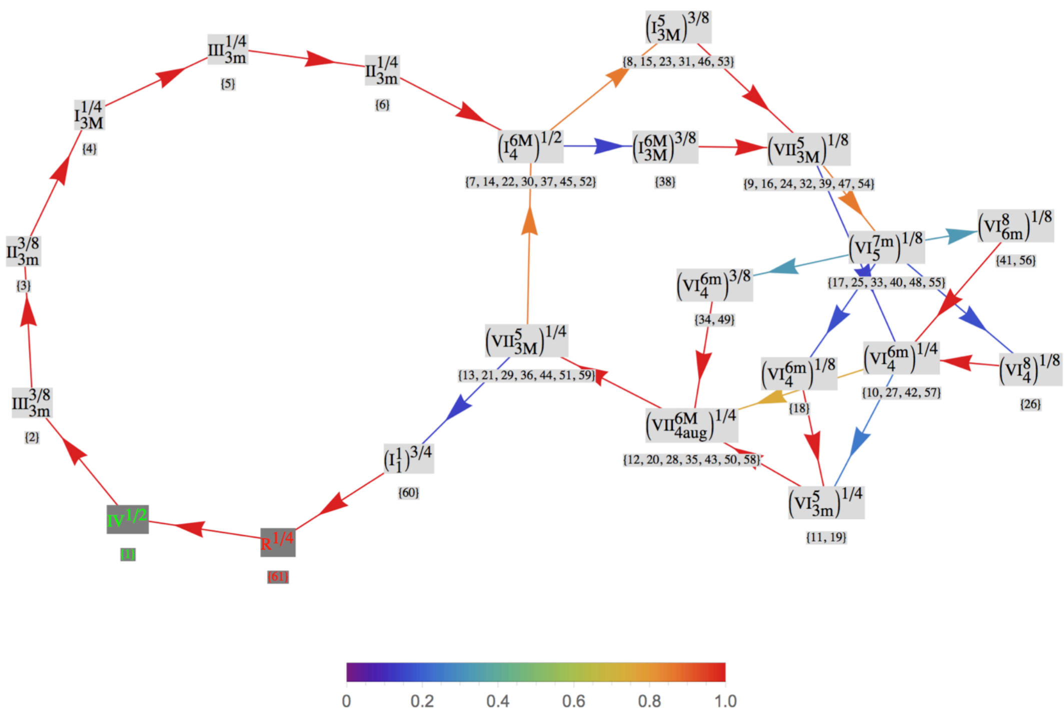

Figure 6. Markov model process graph (right panel) of the complete song *Lushnu Läshgaru*. The chord progression of the actual song (expressed in mode degree notation) can be reconstructed from the sequence indexes below the graph´s nodes.

For the structural analysis of the chord progressions, we removed the introduction, which is typical for many Georgian folk songs (c.f. Fig. 1), and extracted the three-voice part of the song. This results in a considerably simpler process graph (Fig. 7).

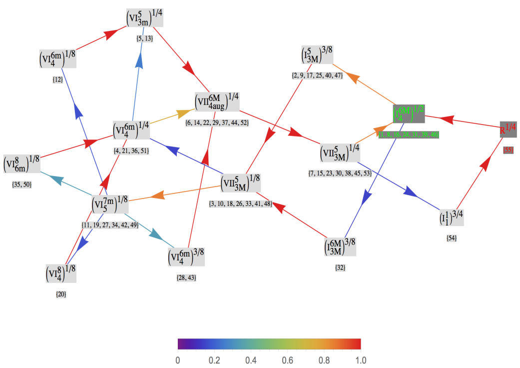

Figure 6. Markov model process graph of the three-voice part of *Lushnu Läshgaru*.

Subsequently we also removed the chord ciphers to obtain the „chord progression skeleton" of the song (Fig. 8), which turns out to be a convenient tool for the structural analysis. In the present case for example it shows that the „building block" of the song consists of the sequence of mode degrees I-VII-VI-VII which is repeated seven times.

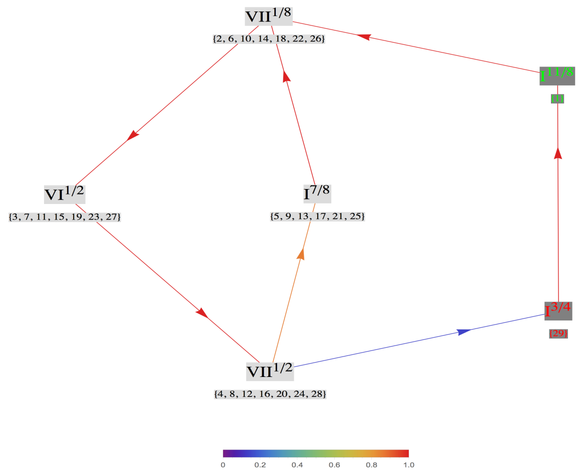

Figure 7. „Chord progression skeleton" of *Lushnu Läshgaru*.

In addition to the process graphs, we spot-checked the quality of individual Markov model representations by listening to random realizations (new synthetic songs) of the models generated from the estimated transition matrices. In general, provided the songs are long enough to allow for a robust statistical analysis of states, the Markov model audibly seems to capture the main characteristics of a song in terms of chord progressions rather well.

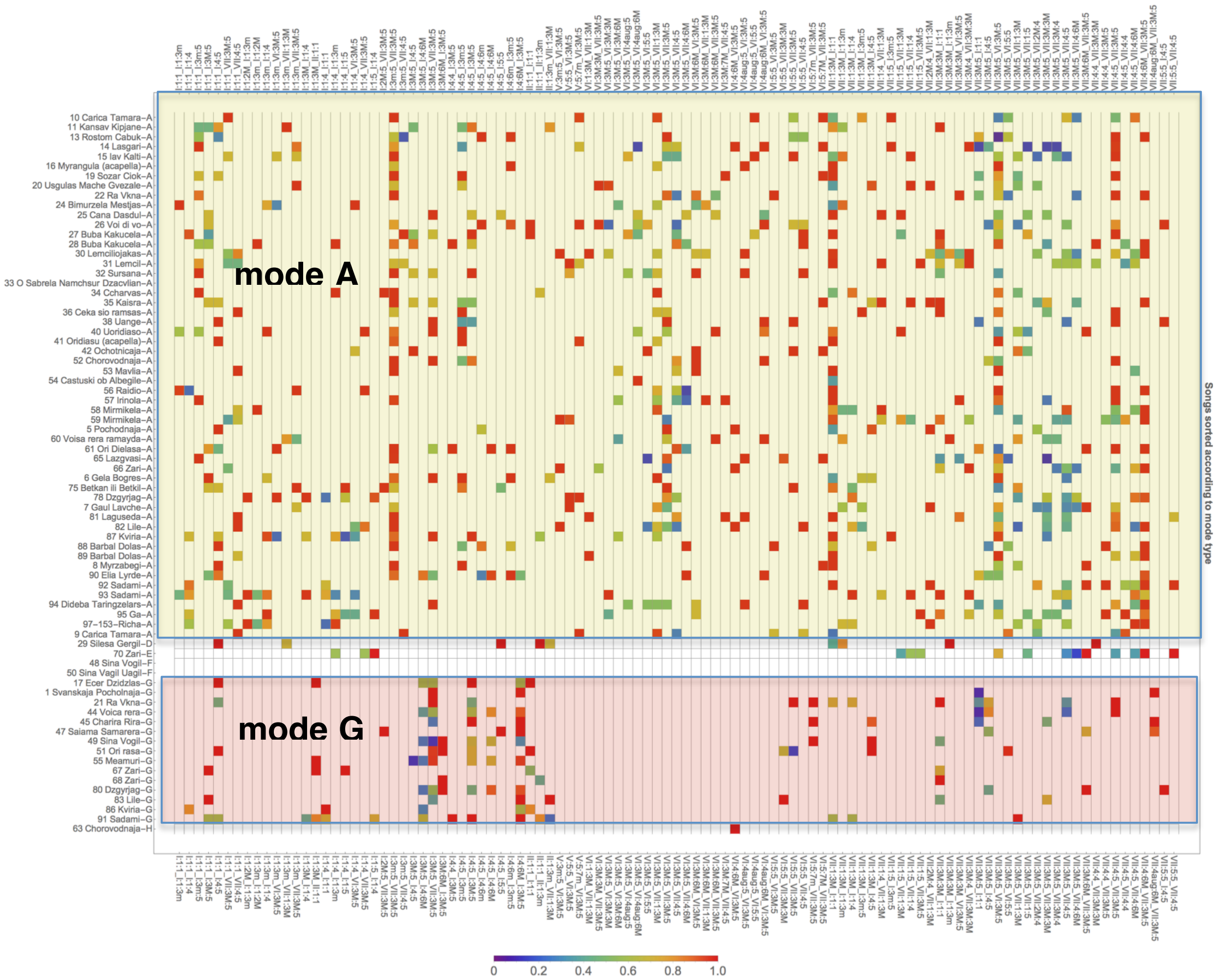

Figure 9. Chord progression profiles regarding the 100 most frequent chord progressions in the whole corpus. The color code of each cell corresponds to the relative frequency with with this chord progression takes place in the corresponding song. Labels are in mode degrees.

4. RESULTS

The state transition matrix T of an individual song (e. g. Fig. 5) can be seen as a simple list of numbers which describe the transition probabilities of all possible chord progressions in that particular song. Usually, the majority of these numbers will be zero because only a small subset of the possible state changes in a song actually takes place. Similarly, for each song one can also determine the transition probabilities of all chord progressions possible in the whole corpus, simply by adding further zeros for all combinations of chords which are in the corpus but not in the particular song. We call this vector the chord-progression vector or -profile of the song. The advantage of doing so is that the chord progression profiles in all songs in the corpus are now of the same dimension and can be quantitatively compared easily. One way of doing so is by projecting all of them into subspaces which are small enough to be visualized (Fig. 9).

Each row in the grid in Fig. 9 corresponds to one song while each column corresponds to one of the 100 most frequent chord progressions in the whole corpus. The songs are sorted vertically according to their modes. The majority of the songs are in mode A (upper yellow shaded rows) and mode G (lower red shaded rows). Those columns which stick out visually as vertical structures correspond to chord-progressions common to all songs of that mode category and are labeled in mode degree notation at the top and the bottom of the panel.

Representing songs in terms of feature vectors (both chord profiles and chord-progression profiles) enables the application of powerful tools from the fields of machine learning and high dimensional visualization to extract information from musical scores. As another example in that direction, Fig. 10 shows the so-called Sammon´s map (Sammon, 1969), which displays the mutual distances of the chord-progression profiles shown in Fig. 9 in such a way that all mutual distances in Fig. 10 are reflecting the mutual distances of the chord-progression profiles in the high-dimensional space.

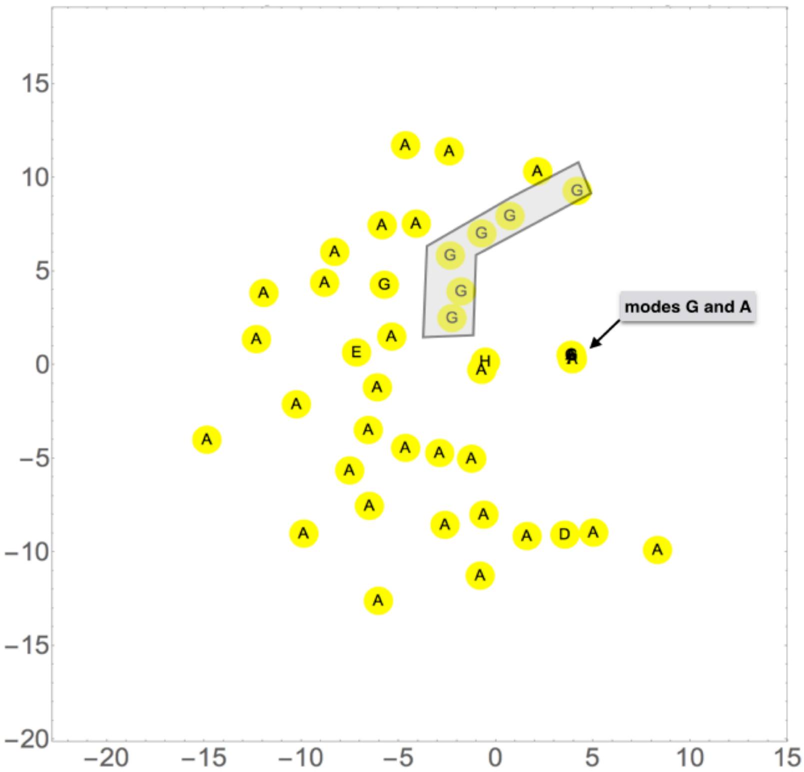

Figure 80. Sammon´s map of the chord-progression profiles of Fig. 9.

Fig. 10 can be directly interpreted as a map of the similarity of the songs in terms of their chord-progressions. Songs which plot close to each other will have more similar chord-progressions than those projecting farther apart. Even without further analysis, Fig. 10 already visually indicates that some of the characteristic chord progressions are mode related.

5. CONCLUSIONS

The present study illustrates different ways to computationally extract information from digital musical scores of polyphonic vocal music and to compare songs quantitatively in terms of chord- and chord-progression structure. The results suggest that Markov models in conjunction with graph theory and high-dimensional visualization techniques provide a powerful and principled framework to perform analysis of musical scores regarding their vertical organization (e. g. Arom and Vallejo, 208; 2010) in semi-automated and completely reproducible ways.

6. ACKNOWLEDGMENT

F.S. is grateful to Giulio Alessandrini from Wolfram Research for fruitful discussions and for sharing his EWTC 2014 presentation on sound processing which stimulated part of the present analysis. We also thank Thoralf Dietrich for his help with the transcription of the scores.

7. REFERENCES

Akhobadze, Vladimer. (1957). *Collection of Georgian (Svan) Folk Songs*. Tbilisi: Shroma da Teknika. (In Georgian. Foreward in Georgian and Russian).

Arom, S., & Vallejo, P. (2008). Towards a theory of the chord syntax of Georgian Polyphony. In Proc. of the 4th International Symposium on Traditional Polyphony, International Research Centre for Traditional Polyphony/ Tbilisi State Conservatory, pp. 321-335, Tbilisi.

Arom, S., & Vallejo, P. (2010). Outline of a syntax of chords in some songs from Samegrelo. In Proc. of the 5th International Symposium on Traditional Polyphony, International Research Centre for Traditional Polyphony/ Tbilisi State Conservatory, pp. 266–277, Tbilisi.

Meyn, S. P. & Tweedie, R. L. (2005). Markov chains and stochastic stability, pp. 567, Springer.

Sammon, J. W. (1969). A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, *C-18*(5), 401–409.

Wolfram Research Inc. (2014). Mathematica, Version 10.0, Champaign, IL, USA.