On pitch and chord stability in folk song variation retrieval

Jörg Garbers, Anja Volk, Peter van Kranenburg,
Frans Wiering, Louis Grijp, Remco C. Veltkamp
Utrecht University
Meertens Institute

Abstract

In this paper we develop a set of methods for computer aided folk song variation research. We examine notions and examples of stability for pitches and implied chords for a group of melodic variations. To do this we study visualizations of extracted features of the aligned melodies. We present how we can use the extractions to find other good candidates for that melody group by redefining the ground distance in our melodic search engine.

1 Introduction

The goal of the WITCHCRAFT project (*What Is Topical in Cultural Heritage: Content-based Retrieval Among Folksong Tunes*) is to develop a content-based retrieval system for a large number of folk song melodies stored as audio and notation. Its purpose is on the one hand to aid folk song researchers in tracing and classifying variations of volk songs and on the other hand to allow the general public to search for melodies with a simple *Query by Humming* or *Keyboard* interface.

Representing melodies and melodic queries as weighted point sets in the onset-pitch domain, as done in the Muugle system [1], proved to perform well in combination with a couple of pre- and postprocessing methods in the general public query task. [7] In the initial part of our project we have tested Muugle's fitness on a test corpus of 141 symbolically encoded Dutch folk songs for the purpose of the other task: variation classification. Although the results were quite promising, it became clear, that a system, that uses more information from the user-query, from the data and from additional feature extractors, can enable researchers to retrieve and classify folk songs in more informed ways. This paper is about this topic.

Overview

We assume that classification and retrieval of melodic variations can benefit from the investigation in *stable features* accross known sets of related music. To know what typically remains stable from one variation of a melody to the next allows folk song researchers to decide, if any given melody belongs to a variation group or not.

To support this kind of classification with a search engine, we must find ways to formulate queries, that specify what should be matched and how strictly. For convenience we prefer to automatically derive such queries from a set of melodies, that are known to belong to the variation group.

Note that the examples and figures, included in this article, are only given to exemplify our computer aided methods. All figures (except for the first) were automatically generated from Humdrum-`**kern` sources with the help of the Humdrum [5] toolkit, the Guido noteviewer [3], Rubato [6] and additional scripts that are to be executed once.

In section 2 we develop modifications to our computational framework to allow to query for similar variants of a melody given a set of examples. In section 3 we examine the stability of pitch across variants and how to use this information for querying and in section 4 we do the same for chords that can be implied from the variations.

2 Retrieval System Modifications

The Muugle system compares a melody query, given as a sequence of events in the onset-pitch-duration domain, with melodies from a database and comes up with a ranked list of close matches. It does so (in principle) by computing the so called "Earth-Movers-Distance" (EMD) between the query melody and any database melody, represented as weighted point sets. (See other paper by Anja Volk et. al. in this volume.)

By representing the note durations as weights, Muugle assures that it allways matches similar amounts of musical duration. Besides of this the EMD requires the definition of a "ground distance", which is defined in the Muugle system as the euclidean distance in the onset-pitch domain, with pitches given in semitones and onset in seconds. A scaling factor in the onset dimension is used to balance the influence of pitch and timing differences in computing an overall distance.

For the following sections we need a generalization or our 'melody x melody' matching approach. There we want to match a pitch q against a set or distribution of alternative pitches P for the same onset. We will use event sequences of such alternatives as queries to find matching melodies in the database.

To formulate these kinds of query within the EMD, we simply have to redefine the pitch distance component of the ground distance in two ways:

a) Let P be a *set* of (alternative) pitches and q a fixed pitch. Then the *minimum* pitch distance between q and P is the minimum of the distances between q and any of the pitches in P.

b) Let P be a *distribution* of pitches and q a fixed pitch. Then the *average* distance between q and P is the weighted average of the distances between q and the pitches in P.

We leave the pre- and postprocessing features of the Muugle system intact, that allow us to compute and combine partial matches and gain transposition and tempo invariance.

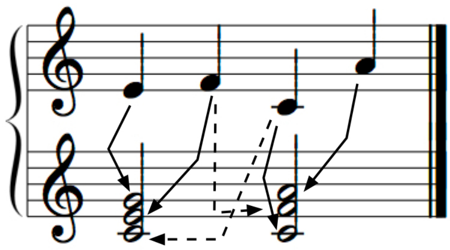

Figure 1: Two EMD matching options for F and C with the refined pitch distance a).

Figure 1 illustrates the effects of the refined ground distance (a). All note duration must flow from the melody to the chords. The E clearly flows to the C major chord and the A to the F major chord. The rest depends on the onset scaling factor: If it is large then F selects a ‘close mismatch’ both in time and pich and matches with the remaining duration from C major chord and C matches perfectly. If it is low, then F matches with the F from F major chord and the C satisfies the remaining duration of the C major chord.

3 Pitch Stability

In this section we develop methods that help to investigate the pitch variability of a given group of melodic variants and that help to automatically find new good candidates for such a group.

3.1 Metrical Levels

Metrical information, such as time signatures (4/4, 6/8) and barlines in common music notation, is used to encode note accent regularities on the time axis. As a working hypothesis we assume that metrically more accented notes are more *stable* across folk song variations than less accented notes. That means that we expect smaller pitch distributions on accented onsets when looking at different variations at the same time.

We try to prove this hypothesis qualtitatively by using the Humdrum *metpos* command to mark the notes of each folk song in a related set with its position within a metric hierarchy. (Levels: 1=bar, 2=half-bar, 3=eights, 4=sixteenth) Then we align the songs in each set by droping upbeats and unmatched verses. For each metrical level we extract all notes above that level and produce views to compare the projection behaviours on the different levels.

When looking for the characteristics of a melody group, we start methodically with very abstract views and proceed to detail views, if necessary. Figures 2-4 show some automatically derived views for the manually aligned variation group 'Frankrijk B1' of *Onder de groene Linde*. [2]

3.2 Evaluation of Pitch Stability

Figure 3 gives us a quick view on the pitch material used per bar at the different metrical levels. By definition we get less or an equal number of pitches at higher metrical levels. But it is interesting to see that there are quite different ranges, both in pitch number and ambitus: While the variation in pitch in bar 3 is reflected on all metrical levels, the variation vanishes on higher metrical levels in bar 7. This might lead us to different matching strategies for different segments (e.g. contour search vs. chord search) when looking for additional members of this variation family in the database.

Figure 4 provides us with a slightly more detailed snapshots accross all variants taken at different metrically motivated grid positions. The note stability increases from the sixteenth up to the half-bar-level but not up to the bar level. In bar 6, second beat, we have even more stability than on the first beat.

To investigate further, where the remaining instability comes from, we look at figure 2 and check the onset positions where the bottom staff contains more than one note. By looking at the other staffs we check, how many variants are responsible for each pitch. In some cases all variants agree except for one outlier (often the first line, e.g. in bar 4), in other cases we find corresponding subgroups within the variants (e.g. last beat in bar 4). This may lead us to the refinement of variation categories.

3.3 Query Formulation

Assume that we have still unclassified variations in the database. To find other candidates for the variation group under investigation in a large database we can proceed as follows: For the set of given melodies, we compute the pitch sets or pitch distributions for each onset at any metrical level. We construct a query with all alternative pitches or pitch distributions for any onset. We use this for searching in the Muugle database, as described in section 2.

Before actually querying the database, one might also want to refine the query (i.e. the pitch distributions) by hand, to get closer to *the melodic model* that one believes the variations are variants of. For this one could use harmonic information (see next section). Matching for example any (new) melody candidate's first and last bar against the G major chord, seems to be a good generalization (see figure 4).

4 Implied Chord Stability

In this section we develop methods that help to investigate the harmonic variability of a given group of melodic variants and that help to automatically find new good candidates for such a group.

4.1 Harmonization

Not all melodies follow harmonic building principles or have implied harmonizations. However, many melodies do allow genre specific harmonizations or already follow such harmonic constraints. This allows even non-trained musicians to sing an additional voice, in folk songs typically one third or fourth apart.

While melody proceeds at beat tempo or faster, harmony typically changes more slowly at bar or half-bar tempo. Melodies that strongly suggest specific harmonizations often contain chord notes as long notes or on metrically strong beat positions, and using melodic filling tones such as passing tones in a less exposed positions.

When translating actual notes to triads, we enter the domain of interpretation and ambiguity. We interpret the given tones in the light of a *harmonic model* to aid the understanding of the music or to generate accompaniments. In our evaluation we follow the approach described in [4] because we have a harmonic analyzer (HarmoRubette), that produces harmonic information for the *best harmonic path*, available as a tool within Rubato.

4.2 Evaluation of Implied Chord Stability

We have tried different harmonic analytic models, i.e. music theoretical parametrizations of the HarmoRubette. But since they were still far from optimal we do not go into their details, but show the preliminary results for the different metrical levels.

The HarmoRubette generates for each onset a function symbol and a key, such as `S(G)`, the subdominant in G. When running the automatic analysis on the differently abstracted melodies, we find irritating results. For some variations the results are completely in G major, for some to in C major. That makes T(C) and S(G) look quite dissimilar and symbol sequences difficult to compare.

An option to cope with this is to insist in a common key G or put in more harmonic model knowledge and reevaluate the melodies. We did not follow this approach yet.

Another option is to listen to the represented chords or to compare chord roots instead of the function symbols. We get more symbol stability, but not surprisingly still different interpretations for 'A' pitches in G and in C.

Already now, the resulting set of chords and – in this case – the extended set of derivable chord notes can be used as note distributions in the extended distance measure of section 2. However, this will only sort out harmonically distanct melodies but will not result in a fine ranking, because it allows very many melodic alternatives.

4.3 Contextualization

The HarmoRubette chose different tonalities, because it analyzed very underspecified chord sequences, where each chord consisted of just one melody tone. When doing so, the analytic ambiguity is inherent. What we prefer is to analyze more constraining chord sequences. In the following we present several ideas where these more constraining additional notes may come from in the environment.

First, the additional notes can come from notes on lower metrical levels following a beat. Figure 2 shows many examples, where often from three 8th notes two can be considered chord tones. In practice the harmonic analyzer can be left alone to figure out, which notes make sense as chord notes in the larger context. We simply have to feed it all notes at once.

Another option is to try to derive a common chord scheme from the whole set of variations. We tested this by running the analysis on the chords of figure 4 and found the consistent key G major with some very short deviations to 'ii' and 'vi' on lower metrical levels. (See figure 5.)

From this we might conclude, that to test, if a new candidate melody belongs to the variation group, we just have to build the common chord set and see, if the new notes do not 'disturb' the analysis. The HarmoRubette also comes up with note weights that express the conformance of the notes with the analyzed harmonic loci.

However, this might but not be possible in general for manually unprepared melodies, that can differ in slightly shifted onsets. For this the note distribution matching strategy within the EMD seems to be more promising.

5 Summary

For a group of folk song variations we have looked into the note stability and the stability of the 'best harmonic symbol sequence' on the onset, tactus and bar levels. Therefore we developed a set of tools and views that allow us to get a quick impression about the stability of features for a set of variations at different metrical levels. We found them quite useful to visually examine the pitch stability and found our hypothesis verified, that melody tones at strong positions are more stable in variations.

We have also presented the idea to use this information in a refined transportation distance measure, that can match pitch distributions with pitches. Depending on the task at hand, this can be used in combination with other distance measures to perform ranking and filtering tasks. We will further study this methodology and concrete examples within the WITCHCRAFT project. We will also consider to use these insights to improve our general public search engine.

6 Acknowledgements

This work was kindly supported by the Netherlands Organization for Scientific Research within the WITCHCRAFT project NWO 640-003-501, which is part of the program *Continuous Access to Cultural Heritage*. Further we want to thank the developers of the Humdrum and Rubato toolkits and the encoders of the Dutch folk songs, who made this investigation possible.

References

[1] Bosma, M., Veltkamp, R. C., Wiering, F. ”Muugle: A framework for the comparison of Music Information Retrieval methods”. In *Proceedings of the ICMPC 2006*, pp. 1297-1303.

[2] *Onder de groene linde. Verhalende liederen uit de mondelinge overlevering*. Uitgeverij Uniepers, Amsterdam, 1987-1991

[3] Guido: http://guidolib.sourceforge.net/

[4] Garbers, J. ; Noll, T.: New Perpectives of the HarmoRubette. In: *Lluis-Puebla, Emilio, Guerino Mazzola und Thomas Noll (eds.): Perspectives in Mathematical and Computer-Aided Music Theory, Verlag epOs-Music, Osnabrück (2003)*

[5] Humdrum:
http://music-cog.ohio-state.edu/Humdrum/

[6] Rubato: http://www.rubato.org/

[7] Typke, R. *Music Retrieval Based on Melodic Similarity*. PhD thesis, Utrecht University, 2007.

Figure 2: The melodies at metrical levels 3 (eights) and above and chords resulting from projecting all notes. (Bars 3-5)

Figure 3: A collected view of all pitches per bar from figure 4. The staffs refer to reductions to different metrical grids.

Figure 4: Four views of all notes of all variations. Each staff shows the notes that fall on the grid of a particular metrical level.

```
=1             =1             =1
4.g            I(G)           G
4.dd           I(G)           G
=2             =2             =2
4.a            V(G)           D
4.g            I(G)           G
=3             =3             =3
4.b 4.cc       IV(G)          C
4.ee 4.cc 4.dd -(G)           G
=4             =4             =4
4.dd 4.a       V(G)           D
4.dd 4.a       V(G)           D
=5             =5             =5
4.dd           I(G)           G
4.ee           IV(G)          C
=6             =6             =6
4.cc 4.ee      IV(G)          C
4.a            V(G)           D
=7             =7             =7
4.ee 4.dd      I(G)           G
4.dd 4.ee      I(G)           G
=8             =8             =8
4.b 4.g        I(G)           G
4.b 4.g        I(G)           G
==             ==             ==
```

Figure 5: Automatic root analysis of the sequence of alternative pitches at metrical level 2 (half-bar). The left column shows the pitches, the right columns show their shared functional analysis and root chords.