Bridging Music Information Retrieval and Folk Song Research
The computational setup of the WITCHCRAFT project

J. Garbers
Utrecht University, The Netherlands, Email: garbers@cs.uu.nl

Abstract

The WITCHCRAFT project sets as its objective to develop a fully functional content-based retrieval system for folk song melodies. Besides making the Netherlands folk song collection of the Meertens Institute available for the general public, we aim to support their research in oral variation. Folk song researchers traditionally know many songs by heart and intuitively find correlations between them. For large or international collections, however, researchers cannot know all the song instances, so computer-aided support is indispensable to reveal related items. In principle this means to take existing models about oral variation in designing and implementing similarity measures that can be used in a retrieval system. This, however, proves to be difficult in practice.

In this paper we present our information retrieval setup consisting of performed folk song recordings, their symbolic encoding, an expert classification of variant tunes (ground truth), annotation data about the particular expert classification reasons and our approaches to similarity. As an example we present how we translated the expert concept of "pitch stability" into an alignment-based similarity measure and how we evaluated our method and the original "pitch stability" hypothesis with respect to the given annotations.

Introduction

In general an information retrieval system consists of a database, a query interface, one or more matching algorithms to compare items from the database with the query and a result presentation component. The database is obviously a limiting factor, because querying for features which are not encoded in the database is meaningless. But the query and response interfaces are also important, because they allow the user to express and satisfy his or her information need. Once the database and the user interface are fixed, the results of different matching methods can be compared with each other.

In the following sections we consider two retrieval system setups, one for the general public and one for folk song researchers. Both are based on the encoded transcriptions of folk songs. We introduce the folk song collection which we make accessible by content-based search. We explain the information need of folk song researchers and sketch options for realizing the folk song query and matching components.

The collection

Onder de groene linde (Under the green lime tree) was the name of a Netherlands radio broadcast show of folk song field recordings from 1957-1994. [1] Old people were asked to sing tunes that they learned in their childhood (mostly by listening). About 5000 of these recordings were manually transcribed by experts before the start of our project (see figure 1). Today all recordings and the scanned transcriptions are part of the *Liederenbank*[1] presentation.

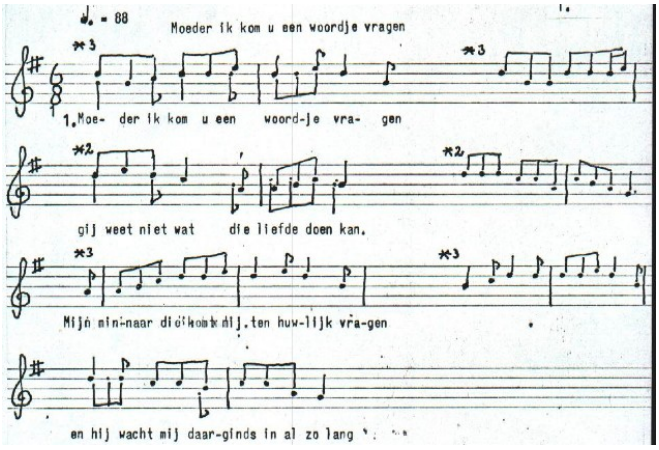

Figure 1: A manual transcription of several strophes of a folk song recording

However, the musical content of the data (both the recordings and the handwritten scores) is difficult to access in a content-based retrieval system. Therefore we developed the necessary editing tools (see figure 2) and guided the processes at the Meertens Institute to encode the transcriptions.

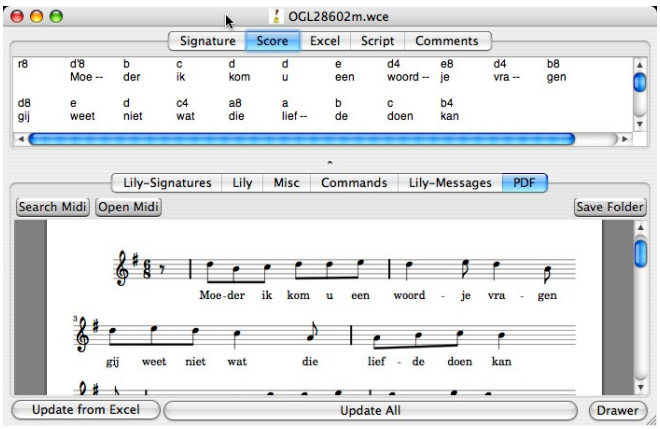

Figure 2: The WitchCraftEditor

[1] http://www.liederenbank.nl

We decided to stay close to the handwritten transcriptions. From these we keep the lyrics-oriented, line-based notation. This results into musically reasonable segmentations, which is handy if users are not interested in matching melody segments that cross line boundaries, or if they want to look only for variants of a particular line of a song. From this encoding we generate representations of the strophes and phrases in MIDI format and in the musicology-friendly Humdrum format.[2] Sheet music images are generated for verification and for the web presentation.

The Liederenbank now presents the first phrase of a song along with metadata and provides a very convenient browse interface: A button allows to search for melodies similar to the current one (see figure 3). A ranked result list is produced using a particular computational model of music similarity. In the future users can choose from several pre-computed similarity relations. However, querying with one example song at a time is not the only option for experts, who should be able to incorporate more analytic and corpus knowledge.

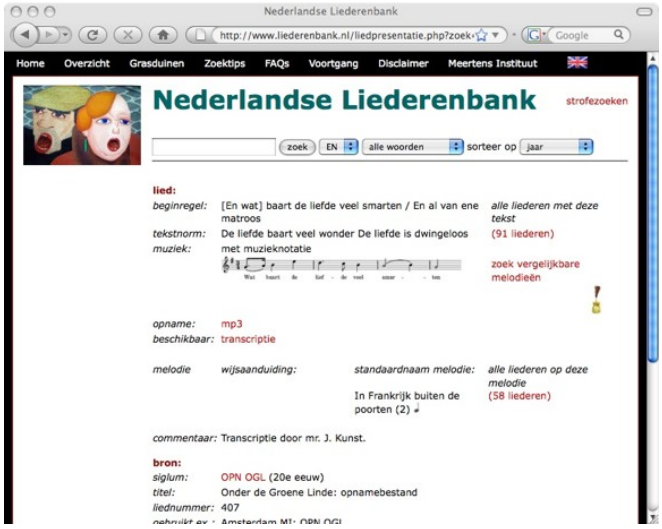

Figure 3: A query-by-example link (*zoek vergelijkbare melodieën*), embedded in the Liederenbank interface

Folk song research models

In [2] we adopted a generic information retrieval system to suit the musical domain. The system allows to instantiate and combine different similarity measures (see figure 4). But to further adopt existing measures for folk song research we must understand the information need of folk song researchers.

Folk song researchers aim to reconstruct the process of oral transmission, where songs are passed from generation to generation by singing, listening and creatively changing them. Songs that are no more part of our culture and that are not preserved by recording or writing them down are lost. Mostly based on profound knowledge and professional intuition folk song researchers relate melodies, that have some identifying

[2] `http://dactyl.som.ohio-state.edu/Humdrum/`

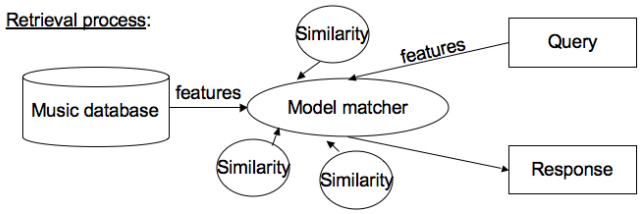

Figure 4: A retrieval system with different similarity measures to choose from

features in common. In the Liederenbank experts assign to melodies such a *melody norm* label. This assignment is not always unproblematic because sometimes songs are composed by putting parts of two different parent songs together. Nevertheless, we can use the melody norm as a *ground truth* for measuring the retrieval performance of similarity measures that we design to find folk song variants.

Melody norm annotations

We asked our folk song researchers to annotate the inter-relations of the melody norm members further for some tune families. After some discussions about which features are relevant and how to encode relations both quickly and with reliable agreement across experts, we implemented an annotation tool (see figure 5). The annotators choose a reference melody for each melody norm and compare it globally and phrase-wise with other members. They rate the song pair's contour similarity and rhythmical similarity and identify common motifs.

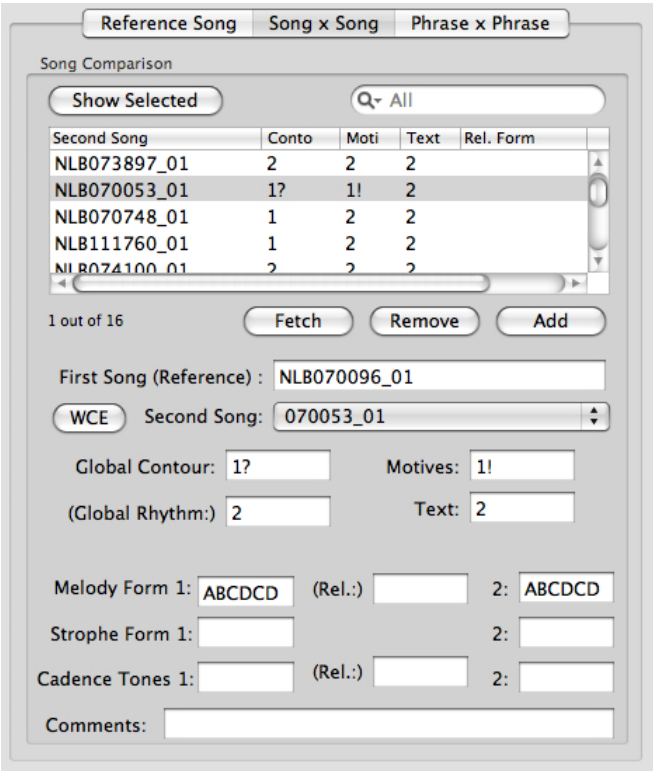

Figure 5: A melody norm annotation tool for global and phrase-wise song comparison

We can use this data in three ways. First, we can

infer which features are important for a specific melody norm and match only those features. Second, we can build evaluation sets that are specific to test contour measures or rhythmical measures. Third, we can align the corresponding phrases (see figure 6) and find what remains stable in a particular group.

70096	Contour	Rhythm	Motives	phrase-1	phrase-2	phrase-3
73277	1	2!	1	1	2	3
71064	2	2	2	1	3	4
70748	1	2	2	1	2	3

Figure 6: Similarity values between reference melody 70096 and other melody norm members and the correspondence of the melodic phrases with those of the reference melody

Pitch stability

We have conducted a *pitch stability* study in [4]. The basic assumption behind the common *pitch stability* hypothesis is that some pitches or pitch progressions are somewhat characteristic for a melody norm. Figure 7 shows a manual alignment of six folk song variants from the same melody norm. Pitches on corresponding onset positions are projected on the staffs below. One can see immediately that there is generally very little variance (usually only two or three pitch classes) and that strong beat positions are more stable that week one's.

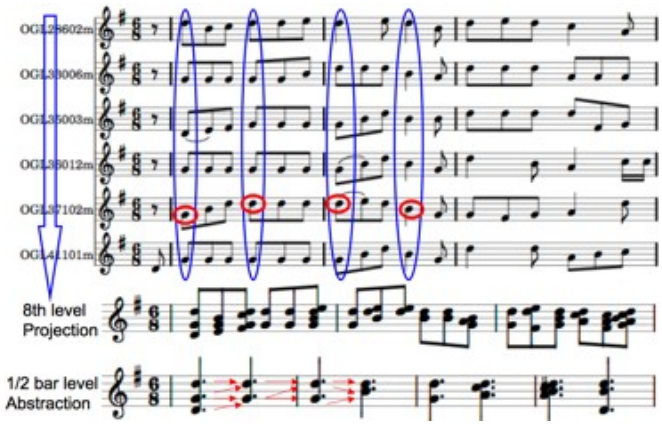

Figure 7: Studying pitch stability using alignments and projections

In [3] we used different superpositions of the aligned variants at different metrical abstractions and found that a matching method based on that superposition performs better than using single melody matching. This should be theoretically the case, when a third variant can be explained by cross-overs of given variants or by a common harmonic progression. Another finding was that pitches are more stable in the beginning of phrases and in the first and last phrase.

Alignment problems

In [3] we describe manual alignments, which can be made using the WitchCraftEditor. Figure 8 shows a manual alignment, which displays the following three *edit operations*. On long notes or upbeats an extra slurred note can be added. Extra notes can be introduced to accommodate extra syllables. The introduction of an invisible rest in the last bar is necessary to align the melodies to a common meter signature. (In the original transcription of the second melody there is a 9/8 measure signature change just for the last bar.)

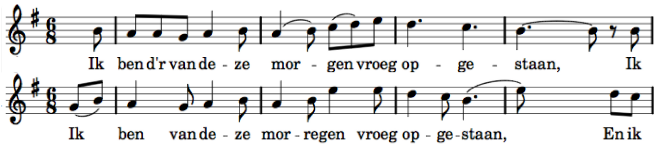

Figure 8: Alignment of 2 variants showing different edit operations

Making manual alignments is quite time consuming. To make querying more practical, we therefore investigate automatic alignment of melody norm melodies. The main problem is to model the type of changes that can occur. [6] In [5] we propose a hierarchically constrained alignment approach which uses segments given by the phrase structure and by the metrical hierarchy. In our current research we study how to best incorporate different musically reasonable *edit operations* into either scoring functions or complexer alignment algorithm setups.

Conclusion

Designing and realizing a music information retrieval system for folk song researchers is quite demanding. We designed processes and tools for data entry and to get information from the researchers. We believe that researchers should be able to pose explicit queries and handle adequately complex user interfaces that enable them to bring in their analytic and intuitive capabilities. Therefore it will probably be more important to support the researchers in formulating good queries than to aim for the ultimate similarity measure.

References

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

[2] J . Garbers. An integrated MIR programming and testing environment. In *Proceedings of the seventh International Conference on Music Information Retrieval.* University of Victoria, 2006.

[3] J. Garbers, P. van Kranenburg, A. Volk, F. Wiering, L. Grijp, and R. C. Veltkamp. Using pitch stability among a group of aligned query melodies to retrieve unidentified variant melodies. In Simon Dixon, David Bainbridge, and Rainer Typke, editors, *Proceedings of the eighth International Conference on Music Information Retrieval*, pages 451–456. Austrian Computer Society, 2007.

[4] J. Garbers, A. Volk, P. van Kranenburg, F. Wiering, L. Grijp, and R. C. Veltkamp. On pitch and chord stability in folk song variation retrieval. In *Proceedings of the First International Conference*

of the Society for Mathematics and Computation in Music (to be published), 2007. (`http://www.mcm2007.info/pdf/fri3a-garbers.pdf`).

[5] J. Garbers and F. Wiering. Towards structural alignment of folk songs. In J. P. Bello and E. Chew, editors, *Proceedings of the nineth International Conference on Music Information Retrieval*, 2008.

[6] M. Mongeau and D. Sankoff. Comparison of musical sequences. In W. Hewlett and E. Selfridge-Field, editors, *Melodic Similarity Concepts, Procedures, and Applications*. MIT Press, Cambridge, 1990. (`http://www.mip.ups-tlse.fr/~mongeau/music.pdf`).