MELOWORKS - AN INTEGRATED RETRIEVAL, ANALYSIS AND E-LEARNING PLATFORM FOR MELODY RESEARCH

Klaus Frieler, Frank Höger, Jörg Korries
Universität Hamburg
Musikwissenschaftliches Institut
Corresponding author:klaus.frieler@uni-hamburg.de

ABSTRACT

In this paper we present the first release of a new integrated melody search and analysis online platform called MELOWORKS. It is built upon a database which incorporates currently about 7,000 tunes from the Essen folk song collection. Songs can be searched and retrieved and statistically analyzed using a modular analysis framework based on symbolical computations.

1. INTRODUCTION

Most likely, techniques of music information retrieval were first enterprised the fields of ethno- and comparative musicology, even as early as 1949 [2], since these discipline had to deal which large amounts of musical data (field recordings, printed folk song collections etc.) Main research objectives have been and still are categorisation and classification of vocal and instrumental folk music with regard to ethnical and geographical origin, functional and melodical types etc., where different feature sets (e.g. phrase and accents structures, tonality, cadences, melodic contour, etc.) are employed. Specific techniques for indexing were developed for content-based retrieval of melodies [1], which are inseparably intermingled with analytical and classificatory reasoning, e. g., by employing the concept of "tune families" based on melodic features [1]. [2] The recently emerging efforts (e. g. [9, 15]) of connecting music information retrieval to folk song research via computational musicology is – in some way – like bringing music information retrieval "back home".

However, to our view there is need for even one more ingredient to a real up-to-date research in this area, namely music psychology. Algorithms of computational muscicology could be build upon and enhanced by results from music psychological research, thereby boosting its ecological validity. In return, we believe that music psychology should be informed and supplemented by statistical investigations of large music corpora, since statistical approaches to music cognition are becoming more and more relevant besides rule-based thinking (cf. [7, 8]). This coincides nicely with the many approaches taken in current MIR employing statistical, particularly Bayesian methods. The success of this methods, e.g., in the MIREX competions, were the evaluation is mostly based on human judgments, hints to an integrated view. Human perception is mostly determined by enculturation built upon innate biological systems, i. e., statistical learning from a music environment under certain constraints. Furthermore, there are a lot of socially determined factors, which to a great extend do not depend directly on the musical objects themselves but on the semiotics of the social and cultural context, e. g., the image of a pop star or the ritualistic use of shamanistic songs and church hymns, which is basically the object of ethnological or sociological studies. The real-world behavior of music users, which is of utmost importance to MIR, is thus determined by psychological (mostly low-level) and socio-ethnological (mostly high-level or extra-musical) factors. To build successful applications, MIR needs to be informed by the fields of ethnomusicology (in the broadest sense), music psychology (including neuromusicology), and music sociology. On the other hand, these fields could greatly benefit from technologies developed in a MIR context, resulting in a classical win-win-situation.

On more point might serve to illustrate this. The ongoing trend of cultural heritage institution to digitise their often huge collections, particularly of field recordings, produces demands and new application areas for MIR techniques, including state-of-the-art audio-based techniques [3]. Ethnomusicology in this regard calls for MIR techniques in an highly integrated manner.

Our contribution MELOWORKS is just a small step on the way to a fully integrated system as envisioned above. As usual, the visions are big, but the money is short. Nevertheless, we already layed the groundwork for a general framework which finally might evolve into a workbench

[1] One of the first example being Ilmari Krohn's classification system for Finnish Folk song from 1903; his system being later on adopted by Bartok and Kodaly [9]

[2] For a comprehensive overview on this topic see [9]. Further examples of early computational projects in ethnomusicology are Alan Lomax' famous Cantrometics projects [10], based on manually extracted feature vectors if folk songs recordings from all over the world, or the automated classification of Slavonic folk tunes with the help of the Wrocław Taxonomy by Anna Czekanowska and co-workers in Poland [4].

[3] Notably, already in 1975 attempts to classify folks songs on the basis of field recordings had been adventured in Scandinavia [13].

for "computational ethnomusicpsychology", and which we are going to describe in more detail in the following sections.

2. GOALS

The MELOWORKS projects aims at an integrated, easily accessible system for retrieval and analysis of monophonic folks songs. It is designed as a flexible, extendable, yet easy-to-use system which can serve for serious scientific research as well as an educational tool for learners and beginners in the field of computational (ethno-)musicology.

3. SYSTEM

A overview of the system is depicted in Figure 1. The core component is a MySQL-database which currently comprises about 7,000 songs from the Essen Collection [12]. The songs are stored as chains of annotated note events along with the original EsAC code and metadata. This double coding of songs allows for most flexible and complete ways of retrieval and analysis. The data was imported via specialized scripts from the original EsAC text files. For the future, it planned that songs can be added to the database via the web interface, which is the main access point to the system [4]. The web interface enables basic retrieval and more advanced analysis operations. Song data and analysis results are displayed together in the browser window for easy comparison. Songs are listenable by a single click on a flash player integrated in the site. Additionally, there are export options for further external use: Songs can be exported as text files in EsAC format and analysis results can be exported as CSV files.

The system is implemented with modern web development techniques. The backend is built with Ruby on Rails and the frontend features state-of-the-art web design and AJAX technologies.

3.1 Retrieval

Selection of melodies is currently only possible on grounds of the historical EsAC song collections (e. g. Kinderlieder, Old-german songs, Ballads, etc.) The user can choose one collection or a single song from a collection at one time. On these selections various analysis procedures can be applied.

3.2 Analysis

The analysis module is based on the conceptual mathematical frameworks of melodic transformations and transformation chains, resp. multiple viewpoints [5]. The basic objects in this framework are sequences of tone events which are represented by onset, pitch and duration. Pitch is represented here as an index in a tone system, which is a necessary generalisation to allow for non-western melodies. Since EsAC does not explicitly specify a tone system, we currently employ the standard MIDI system for indexing

[4] At time of conference MELOWORKS will be available online for the public under http://www.meloworks.uni-hamburg.de.

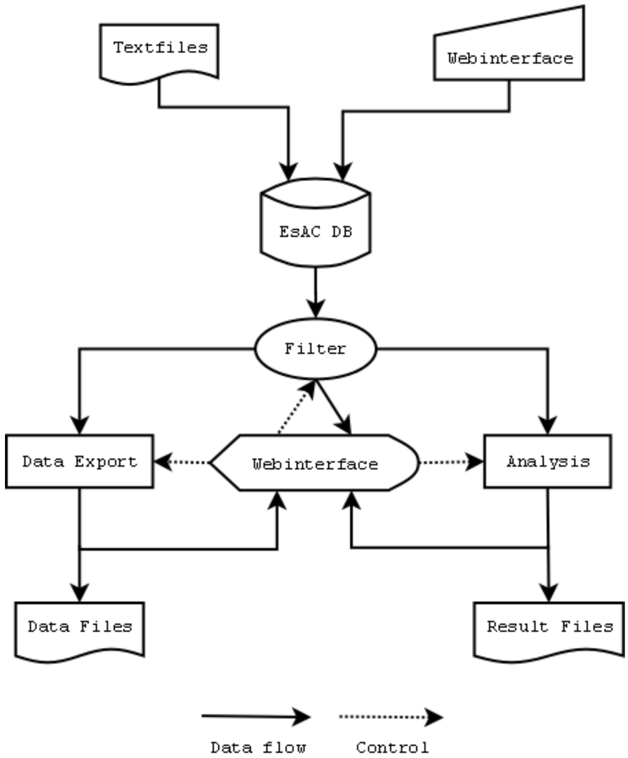

Figure 1. Meloworks system architecture.

tones from the western equal-tempered system. In this framework, a melody of length N is a map

$$\begin{array}{rcl} \mu : [0:N[& \rightarrow & \mathbb{R} \times [0:127[\times \mathbb{R}_0^+ \\ n & \mapsto & (t_n, p_n, d_n) \end{array}$$

where t_n are onsets, p_n denotes MIDI pitches and d_n represent durations. The map must be strictly monotonic in the first argument, i.e., $i < j \Leftrightarrow t_i < t_j$. Furthermore, monophonic melodies must fulfill the condition $t_i + d_i \leq t_{i+1}$ for all $i \in [0 : N-1[$.

For many analytical tasks, it is desirable to measure durations not in seconds (or ticks) but in units of tatums. Since the EsAC durations are alreay denoted using a system of basic rhythmic units (quaver, semi-quaver etc.), we additionally store duration information in the database in integer multiples of basic units. Let $\tau(\mu)$ denote the tatum of a melody, then the unteger durations are calculated as $d'_i = d_i/\tau(\mu) \in \mathbb{N}$.

EsAC provides meter information which is crucial for many analysis applications. Thus, each tone is annotated by a metrical position. Let $\mathcal{B}$ be the coding space of metrical positions consisting of triples (m, b, s) with m denoting the bar index, b the beat index in a bar, and d the subdivision position in a beat (measured in tatums or ticks). Furthermore, let $\mathcal{S}$ be the space of time signatures of pairs (b, t), where b is the number of beats per bar and t the number of tatums (or ticks) per beat. We then define a metrical annotation M for a melody μ of length N as a map $M(\mu) : [0 : N[\rightarrow \mathcal{B} \times \mathcal{S}$. For melodies with changing or free meter (which are quite frequent in folk songs), it is necessary to store a metrical context coded as a timesignature along with the metrical positions [5].

Phrase information is of central importance for folk song research, and its availability is a unique feature of EsAC compared to other symbolical data formats. Hence, phrase information is incorporated by storing phrase indices for every note event.

MELOWORKS currently incorporates 9 melodic transformations which can be combined with n-gram transformations up to $n = 4$. A n-gram transformation for a sequence $\sigma : [0 : N[\to X$ of length N with domain X is a sequence of length $N-n$ with the domain of sequences of length n over X. Clearly, a unigram transformation is the identity map. For example, let $n = 3$, then a trigram transformation of σ has elements

$$\mathbf{3}_\sigma(i) = (\sigma(i), \sigma(i+1), \sigma(i+2)), \quad 0 \le i \le N-3.$$

Let μ be a melody of length N. The 9 melodic transformation currently in MELOWORKS are the following.

1. **Durations**. This transformation maps μ on a sequence of length N of tatum durations $d'_i = d_i/\tau(\mu)$.

2. **Duration ratios**. This transformation maps μ on a sequence of length $N-1$ of duration ratios, i. e. $r_i = \frac{d_{i+1}}{d_i}$. Durations ratios are displayed as fractions of the form "1:1", "1:2", "2:1" etc.

3. **Interonset intervals**. Interonset intervals are defined as the time between two subsequent note onsets. Formally, we map μ on a sequence of length $N-1$ with elements $\Delta t_i = t_{i+1} - t_i$. Interonset intervals are often used instead of durations for analytical purposes.

4. **Semitone intervals**. In the MIDI indexing scheme, pitches are normally assumed the be from the 12-TET system. Thus, the difference of pitch indices is simply the interval between the pitches in semitones. In this case, this transformation is a sequence of length $N-1$ with elements $\Delta p_i = p_{i+1} - p_i$. If in the future the database will contain also music from non-equal-tempered tone systems, a further differentiation between "index differences in a tone system" and "musical intervals" has to be introduced.

5. **Parsons Code**. The Parsons transformation of μ is a sequence of length $N-1$ with elements

$$\mathrm{sgn}(p_{i+1} - p_i),$$

where $\mathrm{sgn}(x)$ is the usual signum-operator. Thus, ascending intervals will be mapped onto $+1$ (displayed as `U`), unisones onto 0 (displayed as `S`) and descending intervals onto -1 (displayed as `D`). Parsons code is sometimes in the literature referred to as melodic contour.

6. **Pitch classes**. This transformation relays on the notion of octave equivalence which allows to separate pitch into the psychophysical dimension of chroma quality and octave position. In MELOWORKS the pitch class transformation uses the key information as provided by EsAC as a reference point. Furthermore, it assumes 12-TET. Let T be a reference pitch, we then define the pitch class of a MIDI pitch p with regard to T in 12-T ET as

$$p_T = [p - T]_{12} = (p - T) \mod 12.$$

Clearly, the tonic and its octavations will be mapped onto 0.

7. **Diatonic pitch classes**. The EsAC code is based on diatonic thinking, i. e., on heptatonic scales. The diatonic pitch class transformation for a given reference pitch T, the tonic of μ, is establish by mapping the pitch classes p_T according to

$$\begin{aligned} 0 &\to 1 \\ 1,2 &\to 2 \\ 3,4 &\to 3 \\ 5,6 &\to 4 \\ 7 &\to 5 \\ 8,9 &\to 6 \\ 10,11 &\to 7 \end{aligned}$$

Strictly speaking, the mapping of the tritone $p_T = 6$ onto the raised fourth is ambiguous, it could be as well conceived as a flatted fifth. However, for folk songs this is not a big issue, since tritones are generally quite rare, and if they occur, then mostly as raised fourths in the role of the leading tone to the dominant.

8. **Tonal pitches**. The tonal pitch transformation with reference tonic $T(\mu)$ is the sequence of translated pitches $p'_i = p_i - T$. According to the EsAC specifiation, the tonic needs to be chosen from the middle octave, i. e., $T \in [60 : 71]$.

9. **Metrical circle map**. The metrical circle map is based on the idea of cyclic time. To this end, every bar of a melody is divided into M equal segments and the metrical positions of note events are mapped onto the according segment index $k \in [0 : M[$. For a more detailed description see [6]. The parameter M can be chosen in the web interface.

This set of transformations is clearly not exhaustive and still under development.

Frequency distributions and statistical values are calculated from the resulting transformations and three different display options can be chosen:

1. The transformation can be viewed directly for each melody;

[5] Of course, certain consistency conditions needs to be fulfilled by such metrical annotations, but we cannot go into detail here.

2. Histograms of frequency distributions are shown for each song. They can be conveniently compared to the corresponding frequency distributions as calculated over the whole song selection;

3. Summaries of numerical statistics can be viewed, again for each song separately and for the whole selection. The statistics comprise number of elements, number of classes, mode, entropy, redundancy [6], as well as arithmetic mean, standard deviation, median, mininum, maximum, and range – if applicable for a transformation.

There is yet another analysis option in MELOWORKS of a slightly different kind, which is dedicated to tonal analysis as given by the Krumhansl-Schmuckler algorithm (see [14] for a discussion of this class of algorithms). For a given selection of songs, the user can chose between six different key profiles that represent weight-vectors for tonalities which are cross-correlated with tone profiles calculated from the songs for each possible transposition. The highest cross-correlation indicates the tonality. All correlations for mode and key are displayed for each song of the current selection. Furthermore, the winning tonality is compared to the tonality as contained in the EsAC data and percentages of correctly classified melodies are calculated. This allows the comparison of different key profiles for different song sets. In a future version, more degrees of freedom of the Krumhansl-Schmuckler algorithm will be controllable, in particular, users will be able to enter their own key profiles.

3.3 Learning

Besides the use for researchers, a platform like MELOWORKS offers excellent learning and teaching opportunities. MELOWORKS lends itself quite naturally for employing the educational paradigm of "explorative learning" [3] in workshop and classes. Students can delve into the principles of ethnomusicology using the different collections by formulating their own research questions and conducting desktop experiments with the online workbench. Furthermore, students are introduced to the basics of computational musicology and thus –ultimatively– to MIR. Last but not least, (statistical) music psychological questions can be explored with the help of MELOWORKS.

We already collected some experience with employing MELOWORKS (and its prototype MELEX [7]) in universitary workshops on melody research and cognition and received promising results and positive feedback from the participants.

4. FUTURE WORK

MELOWORKS, though fully functionable, can be regarded as a prototype for an fully intergrated system as described above. Many extensions and improvements can be thought of.

- **Data**. More data needs to be integrated. In the literature, up to 10,000, sometimes even 14,000 songs are reported for the EsAC collection. We incorporated up to now the songs that are available thru the offical EsAC web site [8], the collections of of Luxembourgian, Lorrainian and Irish Songs by Damien Sagrillo [11], and a set of about 200 songs from Warmia handed down to us by Ewa Dahlig-Turek. A collection of ca. 2,000 Hungarian songs is at our disposal, however with insufficient metadata, hence they are not yet included. Furthermore, we already started collecting transcriptions of monophonic jazz solos coded to EsAC. For the jazz solos we needed to extend the EsAC code and we have to implement a way for dealing with chords as kind of melodic annotations, which is already a first step into polyphony.

- **Retrieval**. One of the most urgent points is the development of retrieval options. These should comprise extensive search by metadata as well as by musical content. The latter will be based on search methods as they are used for example in *Themefinder* [9]. Usability studies could be of great help since designing a satisfactory search interface for melodic content is, in our view, quite challenging. The transformations of the analysis module could be of great help here as well.

- **Analysis**. The analysis module can be greatly extended while sticking to the modular system of transformation chains as far as possible. Music psychologically inspired transformations (e. g., contour, accents) and well-known algorithms from folk song research have highest priority here. Last but not least, a module for similarity calculations is central, which is of particular interest for means of retrieval as well.

  The cherry on the cake would be an easy, but flexible scripting engine, which would enable expert researchers to define their own analysis modules on the base of the existing transformations, display and export options. Yet another idea is implementing an API that would allow accessing the database and the analysis modules remotely.

- **Display & Export**. EsAC code is well readable after short training but more traditional forms of display like common western notation should be provided as well. This can be viewed as part of a more general conversion engine, for exporting melodies in various other symbolical formats (e. g. MIDI, MusicXML, Lilypond).

- **Usability**. A more detailed user rights system could allow to incorporate copyrighted material (e. g., pop

[6] Redundancy is defined here as normalized (inner) entropy substracted from 1

[7] MELEX is based on the same concepts, but written entirely in Javascript with only a few songs incorporated without an underlying database. It is still accessible under http://melex.muo-on.org.

[8] http://www.esac-data.org

[9] http://www.themefinder.org

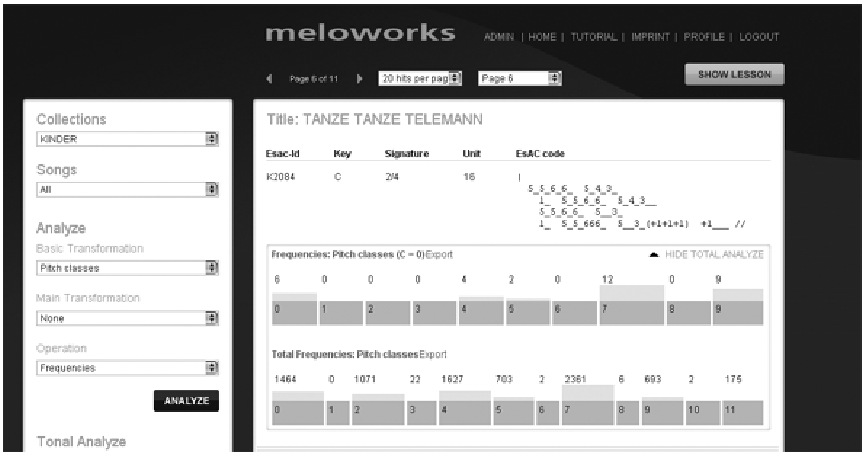

Figure 2. Screenshot of the MELOWORKS web interface.

songs) as well. Users could be given the possibility to save and load retrieval and analysis results for later reuse.

5. CONCLUSION

We presented the first release of an integrated analysis and retrieval workbench for melodies which already has been proven very useful explorative (e-)learning scenarios. The big picture as painted in the introduction is a distant prospect, but a reasonable step in this direction has been done, even the need for further improvements is still huge. We sincerely hope that–after achieving promising first results–we will be able to continue working on MELOWORKS in the future.

Acknowledgments. This work was partly supported by a "Seminare ans Netz" e-learning grant from the University of Hamburg. The web design was carried out by Tino König [10] .

6. REFERENCES

[1] B. Bayard: "Prolegomena to a Study of the Principal Melodic Families of British-American Folk Song," *Journal of American Folklore* , Vol. 63, No. 247, pp. 1–44, 1950.

[2] B. Bronson: "Mechanical Help in Study of Folk Songs," *Journal of American Folklore* , Vol. 62, No. 244, pp. 81–86, 1949.

[3] Bundesassistentenkonferenz: *Forschendes Lernen Wissenschaftliches Prüfen*, Universitäts Verlag Webler, Bielefeld, 2009 (reprint, first issued 1970)

[4] A. Czekanowska: *Ludowe melodie wąskiego zakresu w krajach slowiańskich.*[Narrow-Range Melodies in Slavic Countries], Kraków, 1972

[5] D. Conklin and C. Anagnostopoulou: "Representation and discovery of multiple viewpoint patterns", *Proceedings of the 2001 International Computer Music Conference*, ICMA, San Francisco, 2001

[6] K. Frieler: "Visualizing Music on the Metrical Circle." *Proceedings of the 8th International Symposium on Music Information Retrieval, ISMIR2007*, Wien: OCG, 2007

[7] G. Gigerenzer, D. J. Murray: *Cognition as Intuitive Statistics*, Erlbaum, Hillsdale, NJ, 1987.

[8] D. Huron: *Sweet Anticipation*, MIT Press, Cambridge, MA, 2004.

[9] P. Kranenburg, J. Garbers, A. Volk, F. Wiering, L.P. Grijp, and R. C. Veltkamp: "Collaboration Perspectives for Folk Song Research and Music Information Retrieval: The Indispensable Role of Computational Musicology," *Journal of Interdersiciplinary Music Studies*, doi:10.4407/jima,2009.11.006, 2009.

[10] A. Lomax: *Folk Song Style and Culture*, American Association for the Advancement of Science, Washington, 1968.

[11] D. Sagrillo: *Melodiegestalten im luxemburgischen Volkslied: Zur Anwendung computergestützter Verfahren bei der Klassifikation von Volksliedabschnitten*, Holos, Bonn, 1999

[12] H. Schaffrath: "The ESAC databases and MAPPET software," *Computing in Musicology*, Vol. 8, No. 66, 1992.

[13] R. Sevåg: "Ethnomusicology and Automatic Transcription of Melodies," *NIF Newsletter*, Vol. 2–3, pp. 17–18, 1975.

[14] D. Temperley: *Cognition of Basic Musical Structures*, MIT Press, Cambridge, 2001

[15] G. Tzanetakis, A. Kapur, W. A. Schloss and M. Wright: "Computational Ethnomusicology," *Journal of Interdersiciplinary Music Studies*, Vol. 1, No. (2), 2009.

[10] http://www.macdoct.de