Digital Archive of Finnish Folk Tunes
Application of Musical Data Mining

Music Cognition Group

Petri Toiviainen & Tuomas Eerola
University of Jyväskylä, Department of Music

UNIVERSITY OF JYVÄSKYLÄ

Background

At the end of the 19th century, the Finnish Literary Society (SKS) collected tens of thousands of folk tunes from all over the Finland. Ilmari Krohn published about 9000 of these between the years 1898-1933. This collection is known as the Finnish Folk Songs (Suomen Kansan Sävelmiä). It consists of several subcollections:

- Folk Songs (Laulusävelmät)
- Spiritual Folk Songs (Hengelliset sävelmät)
- Folk Dances (Kansantanssit)
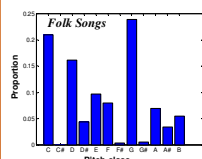- Kantele and Jouhikko Tunes
- Runo-tunes

The collection is musically comprehensive, nationally important and well documented.

Aims

1. Digitizing the Finnish Folk Tunes
2. Building a digital archive
3. Using the archive for research
   - data mining
   - comparative research:
     - musical features
     - lyrics
   - Development of tools for computational analysis (*MidiToolbox*)

Comparison of subcollections

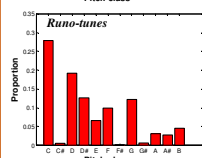*A comparison of musical features of Folk Songs and Runo-tunes*

Pitch-class distributions

Prior to the analysis the tunes have been transposed to either C major or C minor according to the key signature.

Overall, the distributions for both subcollections resemble those observed in other types of tonal music (European Folk Songs, Western Art Music, Bebop Jazz, etc.).

In Runo-tunes the minor third is more common than the major third while in Folk Songs the reverse holds.

A similar difference can be observed in the frequencies of the tonic and the dominant.
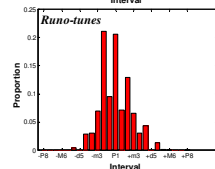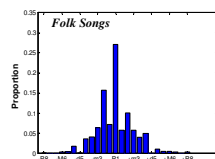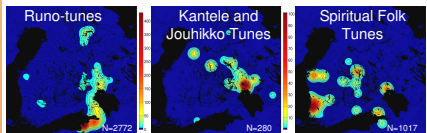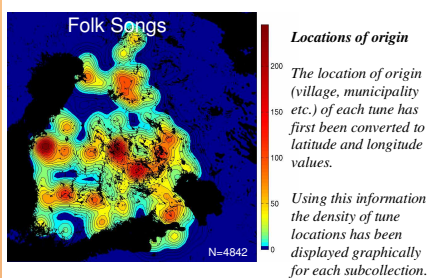
Interval distributions

The interval distributions for both subcollections are typical for Folk music; small intervals are predominant and descending scalar motion more frequent than ascending.

In Folk Songs tone repetition is more common than in Runo-tunes.

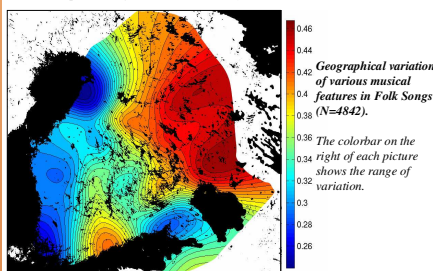Large intervals occur less frequently in Runo-tunes than in Folk Songs.
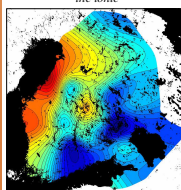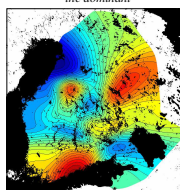

Geographical differences

Folk Songs — N=4842

Locations of origin

The location of origin (village, municipality etc.) of each tune has first been converted to latitude and longitude values.

Using this information the density of tune locations has been displayed graphically for each subcollection.

Runo-tunes N=2772 — Kantele and Jouhikko Tunes N=280 — Spiritual Folk Tunes N=1017

Musical features

Proportion of tunes in minor key

Geographical variation of various musical features in Folk Songs (N=4842).

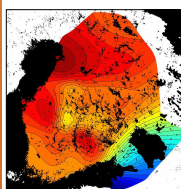*The colorbar on the right of each picture shows the range of variation.*

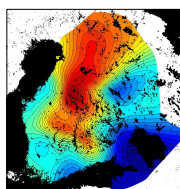*Proportion of tunes starting with the tonic* — *Proportion of tunes starting with the dominant*
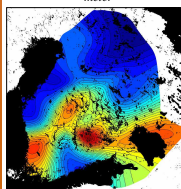

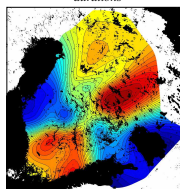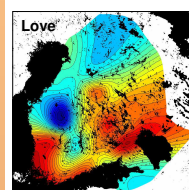*Melodic range (semitones)* — *Average interval size (semitones)*

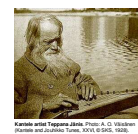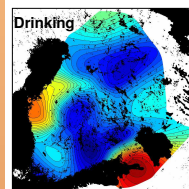*Proportion of tunes in ternary meter* — *Variation (entropy) of note durations*
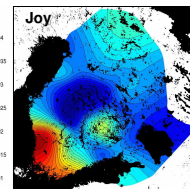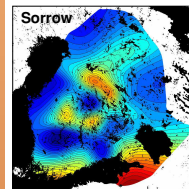

Lyrics

Love

The Folk Song collection contains the lyrics of the first verse of each song. There are a total of 74596 words, of which 13857 are unique (incl. inflectional forms).

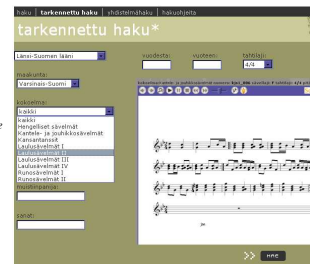Frequencies of words related to some interesting themes are portrayed here.

Drinking

Kantele artist Teppana Jänis. Photo: A. O. Väisänen (Kantele and Jouhikko Tunes, XXVII, © SKS, 1928).

Sorrow — Joy

Search engine

We aim to provide public access to the collection by publishing the archive on the internet. In this application, individual tunes can be searched according to various criteria (villages, regions, lyrics, musical features, collectors). The search engine returns musical notation, MIDI file and textual information for each tune fulfilling the search criteria. We also plan to develop search methods based on musical content (e.g. query-by-whistling).

User-interface of the search engine.

Contact information