



The WITCHCRAFT Project: A Progress Report

Frans Wiering

IMS Study Group Meeting, Zürich, 10 July 2007



Universiteit Utrecht

Talk outline

- CATCH programme
- WITCHCRAFT project
 - aim and team
 - partners and their contribution
 - results and plans
- Beyond WITCHCRAFT
- MIR and Musicology

CATCH

- Continuous Access To Cultural Heritage
 - research programme funded by Netherlands Organisation for Scientific Research (NWO)
- 10 4-year projects
 - Computer Science partner
 - Cultural Heritage partner
 - funding for postdoc, programmer and PhD student
 - employed by CS partner
 - hosted by CH partner
- CATCH projects are (among other things) about
 - automatic annotation and semantic interoperability
 - searching image collections, radio archives and archaeological objects
 - personalised access to collections
- Important aspect: community building and exchange
- CATCH 2 is under development

WITCHCRAFT overview

- partners
 - Meertens Institute (MI)
 - research and documentation of Dutch popular culture
 - folkloristic traditions, multiculturalism
 - popular song from Middle Ages to the present
 - valorisation
 - Department of Information and Computing Sciences, Utrecht University (ICS)
 - broad view of computing
 - content-based Music Information Retrieval (MIR)
- WITCHCRAFT
 - What Is Topical in Cultural Heritage: Content-Based Retrieval Among Folksong Tunes
 - 1 May 2006-1 May 2010
 - aim: to develop a fully-functional content-based retrieval system for large amounts of folksong melodies stored as audio and notation



WITCHCRAFT: the team

- Anja Volk: Postdoc
- Jörg Garbers: Postdoc/Implementation
- Peter van Kranenburg: PhD
- Frans Wiering (project leader)
- Remco Veltkamp (co-leader)
 - head of Multimedia Algorithmics Group, ICS
- Louis Grijp (co-leader)
 - professor of History of Dutch Song, UU and head of DOC Lied, MI
- DOC Lied
 - Martine de Bruin
 - Ellen van der Grijp
 - Ineke van Beersen

IMS Study Group Meeting, Zürich, 10 July 2007

5

Meertens Institute (1): Liederbank



- Database of Dutch Songs
- metadata c. 125.000 songs
- Middle Ages-present
- www.liederenbank.nl
- online since 22-6-2007
 - English interface in September?
- successfull tracing of 'lost melodies' e.g. by means of metrical structure
- *missing: searching musical content*

Waer werd op - rech - ter trouw Dan tus - schen

Oorsponkelijke melodie van de openingsregel van de Rey van Burghzaten uit de Gysbrecht van Aemstel.

Door onze redacteur
KASPER JANSSEN
AMSTERDAM, 7 DEC. Hoe klonk de Gysbrecht van Aemstel 367 jaar geleden? Dat kunnen we weten, nu de muziek van de Rey van Burghzaten uit het toneelstuk van 'Vondel' is teruggevonden. De melodie van 'Waer werd oprecht trouw', gezongen bij de opening van de Amsterdamse Stadschouwburg in 1638, blijkt afkomstig van het lied 'Wanneer elken man en vrouw ter we...

Melodie Gysbrecht gevonden

sten en muziek. In de zestiende en zeventiende eeuw werden die meestal niet samen afgedrukt. Voor veel liederen werd geen nieuwe muziek gecomponeerd. Men zong – zoals nu nog bij bruiloften en feesten – 'op de wijs van'. Die melodische kende liederen. Wanneer de computer constateert dat lengte en ritmiek van woorden en strofen precies passen op de muziek van een lied, kan het om een verloren melodie gaan.

IMS Study Group Meeting, Zürich, 10 July 2007

6

MI (2): Onder de groene linde

- 7.000 field recordings by Ate Doornbosch et al., 1957-1994
- ballads
- last traces of a song tradition that has now almost completely disappeared
- singers: mainly old ladies remembering songs from their youth



IMS Study Group Meeting, Zürich, 10 July 2007

7

Example: Jan Alberts (1)

De Nederlandse Liederenbank

In de Nederlandse Liederenbank zijn meer dan 125.000 Nederlandse liederen online beschikbaar tot de twintigste eeuw. Het gaat om liederlieders, spottlieders, geuzenliederen, psalmen en andere religieuze liederlieders, kinderliederen, orkestratie- en kerstliederen en nog veel meer. Bronnen van de liederen zijn liedboeken, liedbladen, liedkriften en veldwerkopnamen van elk lied is de bron aangegeven waar de tekst en eventueel de melodie is te vinden. Het aantal gezangen kan direct worden doorgeëkt naar de gehele tekst van het lied of naar een gedeelte van de tekst.

De Nederlandse Liederenbank wordt geproduceerd op het Meertens Instituut in het Onderzoekscentrum van het Nederlandse lied in samenwerking met diverse partijkollega's. Het onderzoek wordt voortdurend verdergevoerd aan zowel de presentatie als de inhoud van de Nederlandse Liederenbank.

Wat zit er in de Nederlandse Liederenbank?
Hoe zoek ik in de Nederlandse Liederenbank?

© 2007 XXXX/Meertens Instituut

Nederlandse Liederenbank

zoek | Ent zoeken | alle woorden | sorteert op | jst

23 resultaten

beginregel	afsluiting	melodie	scan
Jan Alberts DeVries/DeVries 1882 (1882), p.119 [nr. 3112]	wijsaanduiding	standaardnummer	melodie
En Jan Alberts stond op en hij zong er een lied	Heer Halbeleg	melodie	scan
OPN OGL 404: opname Houdgehege 1950	Heer Halbeleg	tekst	scan
Jan Alberts stond op en hij zong er een lied	Heer Halbeleg	melodie	scan
OPN OGL 19101: opname Terwisgal 1963	Heer Halbeleg	melodie	scan
Jan Alberts trek uit toch uw bovenste kleed	Heer Halbeleg	melodie	scan
OPN OGL 19102: opname Bellingwilde 1964	Heer Halbeleg	melodie	scan
Jan Alberts stond op en die zong er een lied	Heer Halbeleg	melodie	scan
OPN OGL 19103: opname Drachten 1962	Heer Halbeleg	melodie	scan

http://www.lijderenbank.nl/

IMS Study Group Meeting, Zürich, 10 July 2007

8

Example: Jan Alberts (2)



- Recorded Terwispel, 11-3-1963
- Jan Alberts=Heer Halewijn=Lady Isabel and the Elf-knight
- 115 songs with this theme
- 38 have similar melody
- oral transmission: learned from father
- 'strophic variation'

Jan Alberts stond op en hij zong er een lied
En dit verhoorde een koningskind.

Zij stak er haar hoofd door de venstergaten uit
En toen riep zij: Mag ik er met Jan Alberts uit
rijden gaan?

Jij moogt met Jan Alberts uit rijden wel gaan
Maar als jij er je eretje maar d'r bewaart



IMS Study Group Meeting, Zürich, 10 July 2007

9

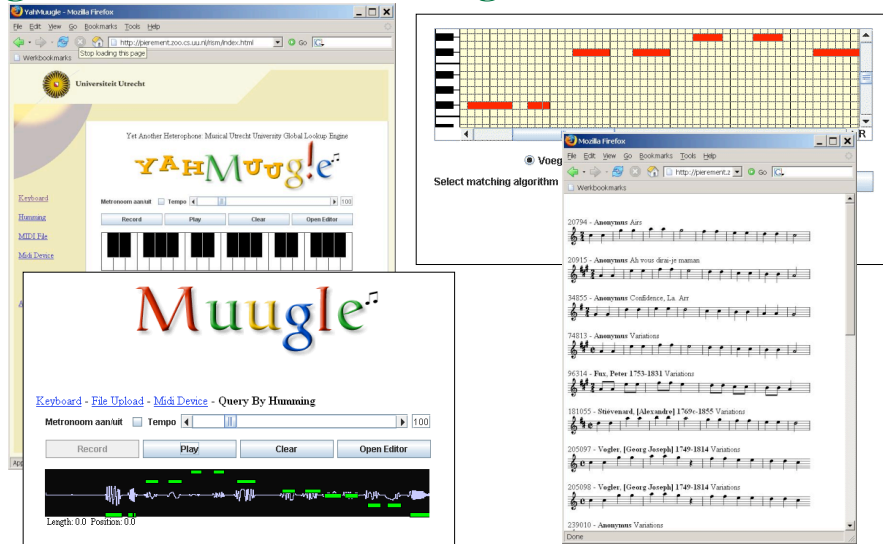
MIR research at the ICS

- fundamental MIR problem:
 - modelling musical similarity
 - algorithmic approach
 - PhD thesis Rainer Typke (Februari 2007): transportation distances
- demonstration/experimentation system: Yahmuugle (Martijn Bosma e.a.)
- C-Minor: Bas de Haas

IMS Study Group Meeting, Zürich, 10 July 2007

10

give-lab.cs.uu.nl/muugle



The image displays the YAHMUUG!e web interface and its associated Muugle application. The top part shows the website with a keyboard interface and a 'Record' button. The bottom part shows the Muugle application with a piano keyboard, a 'Record' button, and a waveform display. To the right, a Mozilla Firefox window shows a list of musical pieces with their corresponding MIDI files and a 'Select matching algorithm' dropdown menu.

IMS Study Group Meeting, Zürich, 10 July 2007

11

Searching the musical content

- motivation: metadata are insufficient
 - no concept of authorship
 - 'contrafacts'
- identification
 - questions from general public
 - supports performers
 - 'melody norm': determine historical relationship between melodies
- **oral transmission**
 - variability of melodies (and texts) related to cognitive processes
 - modelling oral transmission has **added value for each of the partners**
 - MI: modelling results in better understanding of the transmission process
 - UU: better models result in better retrieval

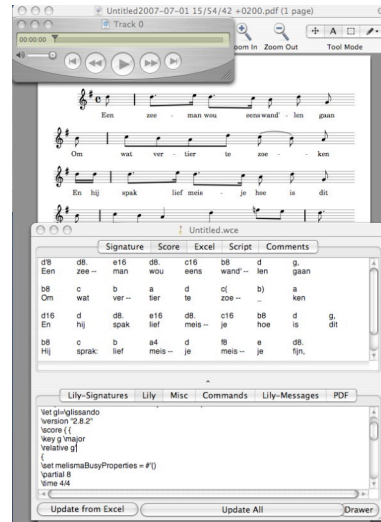


IMS Study Group Meeting, Zürich, 10 July 2007

12

Result 1: corpus building

- OGL songs available as:
 - digital audio
 - manual transcription (scanned)
 - no encoding
- WitchCraft Editor
 - integrates number of tools for Mac
 - simplified Lilypond encoding
 - conversion to MIDI and humdrum
 - preserves verse lines of song
- Test corpus version 1
 - 141 encoded melodies
 - melody groups assigned by specialist
- Dutch Song as Musical Content
 - 2-year project for corpus creation
 - tool for online data entry
 - June 2007: c. 750 melodies encoded by students



IMS Study Group Meeting, Zürich, 10 July 2007

13

Result 2: design/evaluate retrieval methods

- Inner Metric Analysis (Anja Volk)
- Graph matching and indexing (Alberto Pinto, Reinier van Leuken, Fatih Demirci)
- Query-by-group (Jörg Garbers)
- Testing environment (Jörg Garbers)
 - collaboration with Daniel Müllensiefen et al. (Goldsmiths)
 - simplifies experimentation with parameter settings
 - saves intermediate results

IMS Study Group Meeting, Zürich, 10 July 2007

14

Next steps

- **Integration with Liederbank**
 - deadline 1 November 2007 for 1st iteration
 - for internal/expert use only
 - demo of present functionality -> software presentations
- **further evaluation of similarity measures**
 - includes audio (chroma) matching--collaboration with Bonn (Meinard Müller)
- **modelling of oral variation by means of**
 - investigation of stable features
 - expectation-realisation
 - constraints

Ideas for going beyond WITCHCRAFT

- **Collaboration with other popular song/folksong projects**
 - 18th-century French song
 - Hungarian folksong (Zoltán Juhász)
- **RISM**
 - much experimentation done with RISM A/ii
 - time to make 'official' connection
- **Search technology in digital (critical) editions of music**
 - example: Corpus Mensurabilis Musicae Electronicum (Ted Dumitrescu)
- **Access to music archives**
 - several candidates in the Netherlands
- **Creative industry**

MIR and Musicology

- MIR is not computational musicology
 - mismatch MIR and folksong research (Peter van Kranenburg)
 - simplistic ideas about music
 - get better methods in exchange for data
 - ground truth and explanation
 - Willard McCarty: *Humanities Computing* (2006)
 - subtle ideas about interdisciplinarity
 - understanding through modelling
- Can humans benefit from MIR?
 - Information Retrieval and 'aboutness'
 - music as text, creator's perspective
- 'Beyond MIR'
 - music is meaningful--how is meaning generated?
 - MIR as enabling technology for (for example) research and educational applications