

Multimedia-Metadaten und ihre Anwendung

14.02.2006

Inhaltsbasierte Musiksuche
Konzepte und Anwendungen

Dmytro Shaykhit

Einführung

- Die Suche per Text in einer großen Musikbibliothek ist nicht mehr sinnvoll
- *Musik-Content* – beschreibt die Musik selbst (deutsch: Inhalt der Musik)
 - Beispiele: Noten, akustisches Signal, aber auch Text (Lyrik)
- *Musik-Content* als ein weiterer Parameter für die Suche benutzen

Begriffe MIR und MIR-System I

- MIR steht für *music information retrieval*
- MIR nach DFKI (Deutsches Forschungszentrum für Künstliche Intelligenz):

„Musik Information Retrieval stellt ein interdisziplinäres Forschungsgebiet dar, das bereits seit Ende 1960 existiert und sich mit der zunehmenden Verbreitung des Internet und entsprechenden digitaler Formate (bspw. MP3) zunehmend etablieren konnte. Die zu lösenden Kernprobleme umfassen Technologien, die dem Menschen einen effizienten Zugriff auf umfangreiche Musikkollektionen ermöglichen.“

Begriffe MIR und MIR-System II

- MIR-System besteht mindestens aus folgenden Komponenten:
 - Einer oder mehrerer Datenbanken zur Speicherung und Verwaltung der Musikwerken und der entsprechenden Metadaten
 - Einem Server für die Bearbeitung der Benutzeranfragen und die Verwaltung der Datenbanken
 - Einer Anwendung, die die inhaltsbasierte Musiksuche unterstützt
- → Analyse des Musik-Contents ist für eine effiziente Suche notwendig

Musikfeatures

- Musik-Content ist in 7 musikalischen Merkmalen (Features) zerlegbar:
 - Tonhöhe
 - Tempo (Rhythmus)
 - Harmonie (Polyphonie)
 - Klangfarbe
 - Redaktionelle Information
 - Textuelle Information
 - Bibliographische Information

Suchprozess

- Wie funktioniert die Suche nach einem Musikwerk in einem MIR-System?
- Den Suchprozess kann man i.a. in 4 Schritten aufteilen:
 - Eingabe der Musikfeatures des gewünschten Musikstücks
 - Umwandlung der eingegebenen Daten in eine passende Darstellung und Erstellung eines Query-Objektes
 - Matching zwischen dem Query-Objekt und der Information aus der Datenbank
 - Rückgabe einer Liste mit den ähnlichen Musikwerken

Suchprozess: Eingabemöglichkeiten

- Direkte Eingabe der Musikfeatures
 - Beispiel: Tonhöhen-Tempi-Kombination als Suchkriterien in der RISM-Datenbank
$$\%F-4\$bB@3/8\#'8C.6.3\$,B'C\&/,8A'D6$$
$$(-)D/,8G'8.C,6B/8F$$
- Vorsummen (*query by humming*)
- Musikfragmente als Eingabe (*query by example*)
- *Beatboxing* – künstliche Erzeugung der Töne (durch Lippen, Gurgel und Mund)

Suchprozess: Erstellung eines Query-Objektes I

- *Query by humming*
 - Verarbeitung des Vorsummen-Signals für Feststellung der Notenänderung
 - Umwandlung in eine passende Darstellung
 - Mögliche Darstellungen:
 - Beschreibungsschema *MelodyContour* für monophonische Musik in MPEG-7-Audio-Standard
 - Tripel: (Tonhöhenkontur, Tonhöhenintervall, Dauer)

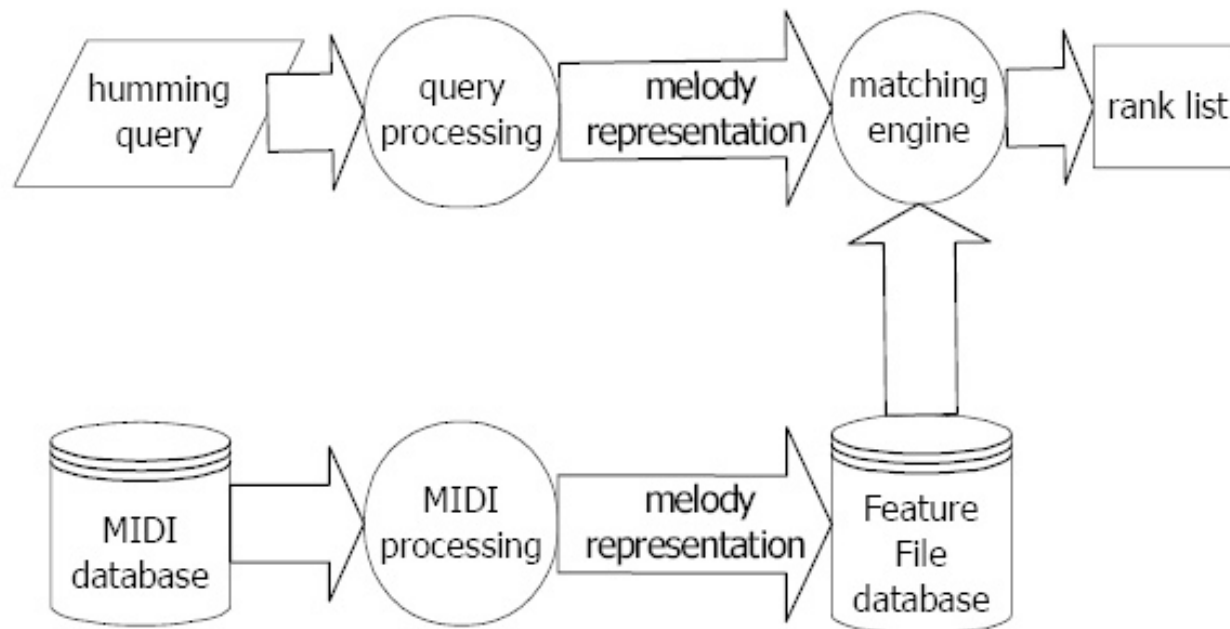
Beispiel:



(* , * , *) (U, 64.1, 1) (U, 71.9, 2) (U, 124.7, 2) (U, 96.0, 3)
(D, -96.0, 1) (D, -124.7, 3) (D, -64.1, 1)

Suchprozess: Erstellung eines Query-Objektes II

– Beispiel eines QBH-Systems:



Quelle: L. Lu, H You, HJ Zhang: A new approach to query by humming in music retrieval. In Proc. of the IEEE ICME01, Tokyo, Japan (2001)

Suchprozess: Erstellung eines Query-Objektes III

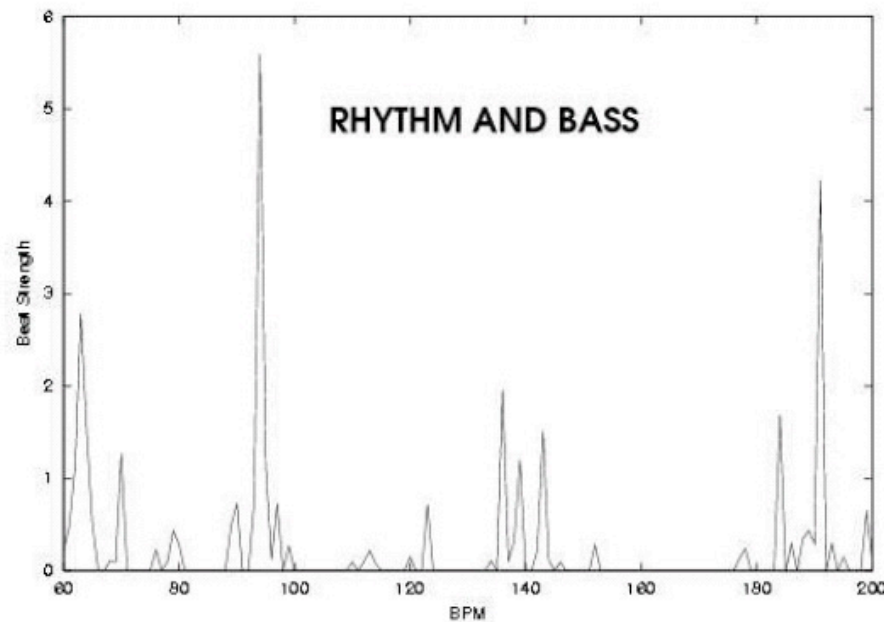
- *Query by example*
 - Ziel: Die Suche nach Musikstücken des gleichen Genres
 - Ein- bis zweisekündiges Musikfragment als Anfrage benutzen
 - Erstellung eines Query-Objektes aus dem Musikfragment (z.B. Extrahieren der Spektralvektoren für polyphonische Musik)

Suchprozess: Erstellung eines Query-Objektes IV

- *Query by beatboxing*
 - Ziel: Die Suche nach *drum loops* (*stark rhythmische Töne*) und auch die Erstellung der neuen *drum loops*
 - Im Gegensatz zu meisten QBH-Anwendungen wird hier nach polyphonische Musik gesucht
 - Berechnung eines Feature-Vektors für die Dauer des Tons

Suchprozess: Erstellung eines Query-Objektes V

- Beispiel für die Darstellung der *drum loops*:
 - *Beat histogram* – zeigt die Verteilung der verschiedenen Taktenperiodizitäten des Signals



Suchprozess: Matching-Algorithmen I

- Zeichenkettenbasierte Methoden für monophonische Musik
 - Musik durch Zeichenketten darstellbar
 - Berechnung der „Stringdistanzen“:
 - Suche nach der größten gemeinsamen Teilsequenz
 - Suche nach der Stellen, wo eine Zeichensequenz in einer anderer Sequenz vorhanden ist
 - Approximiertes Matching: inwiefern die Zeichenketten sich voneinander unterscheiden
 - Indexstrukturen: invertierte Listen, B-Bäume, B*-Bäume

Suchprozess: Matching-Algorithmen II

- Mengenbasierte Methoden
 - Musik als Menge von Ereignissen bzw. Features mit bestimmten Eigenschaften (Tempo, Rhythmus, Tonhöhe)
 - Beispiel: Musik wird durch eine Menge von Feature-Vektoren dargestellt
 - Jeder Vektor wird in einem separaten Array gespeichert
 - Jedes Musikstück ist ein Punkt in \mathbb{R}^n
 - Berechnung der euklidischen Distanz für Musikstücke v_1 und v_2

$$d = \sqrt{\sum (v_1^i - v_2^i)^2}, \quad d \leq \varepsilon$$

Suchprozess: Matching-Algorithmen III

- Musik als Menge von Noten
 - Exaktes Matching: Suche nach der Obermenge des Query-Objekts
 - Approximiertes Matching: Suche nach Obermengen der Untermengen des Query-Objektes
- Indexstrukturen: invertierte Listen

Suchprozess: Matching-Algorithmen IV

- *Probabilistic matching*
 - Untersuchung der probabilistischen Eigenschaften der Kandidaten
 - Beispiel: Berechnung der Markov-Modelle, die die Wahrscheinlichkeit der Zustandsübergänge in Musikstücken beschreiben
 - Zustände entsprechen den Features, wie bestimmte Tonhöhe, Intervall oder Notendauer

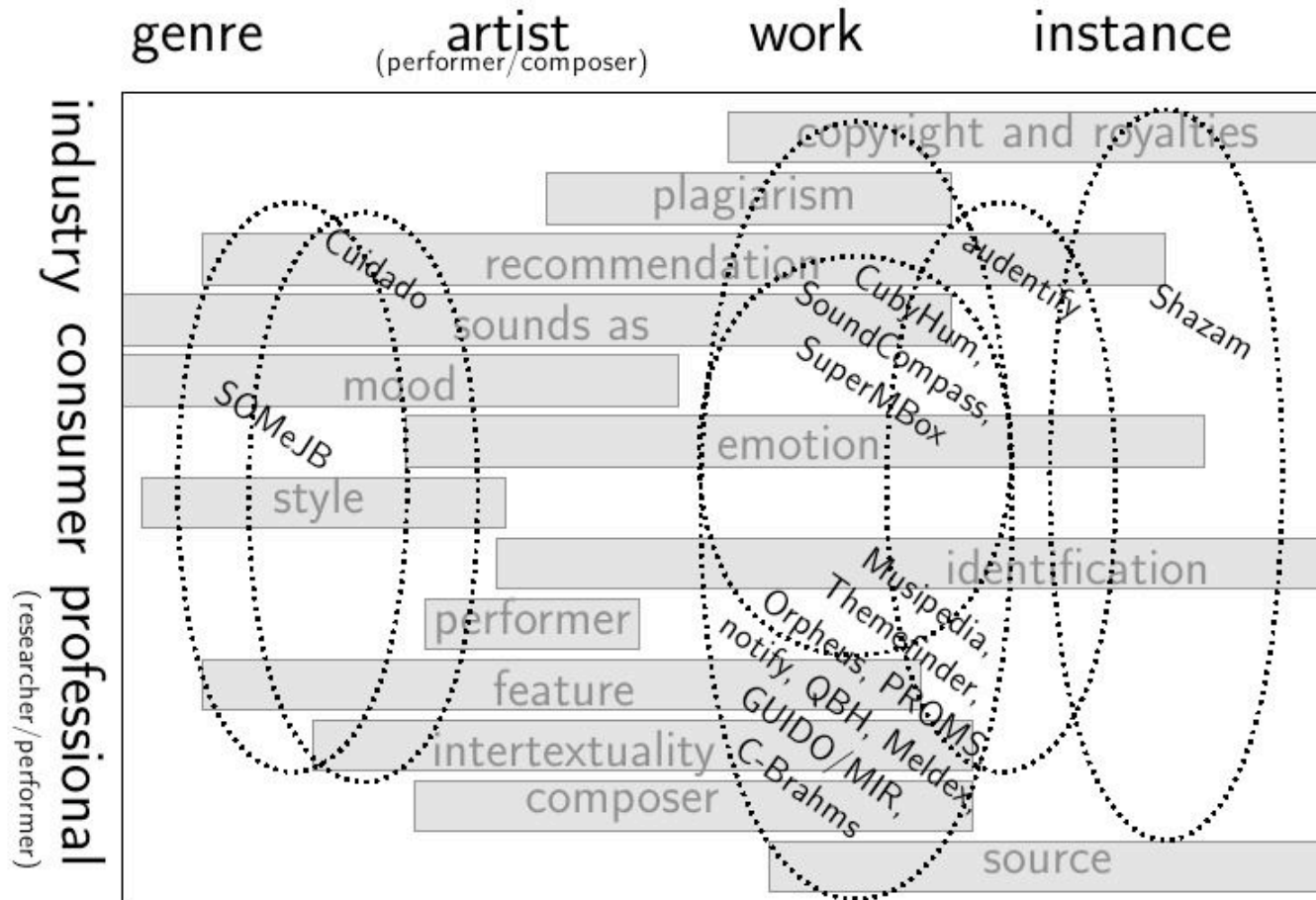
Suchprozess: Matching-Algorithmen V

- Berechnung des Produktes der Wahrscheinlichkeiten basierend auf der Übergangsmatrix des Kandidaten (siehe GUIDO-System: <http://www.informatik.tu-darmstadt.de/AFS/GUIDO/index.html>)
- *Hierarchical clustering*: Speicherung der Übergangsmatrizen von einzelnen Musikstücke in den Blättern eines Baumes.
- Indexstrukturen: Baumstrukturen

MIR-Systeme I

- 3 Benutzergruppen: Industrie, Amateure und Profis (Interpreten, Lehrer, Musiker)
- Aufgabenklassen: Suche nach Genre, Artist, Musikwerk oder bestimmte Instanz des Musikwerks
- Einige der bekanntesten MIR-Systemen:
 - CUIDADO Music Browser: verwaltet die Musik-Meta-Daten in MPEG-7-Format
 - Meldex/Greenstone: speichert monophonische Musik, berechnet Stringdistanzen für Matching
URL: <http://www.nzdl.org/fast-cgibin/music/musiclibrary>
 - GUIDO/MIR: speichert monophonische Musik, verwendet wahrscheinlichkeitsorientierte Methoden (*probabilistic matching*) für Matching
URL: <http://www.informatik.tu-darmstadt.de/AFS/GUIDO/index.html>

MIR-Systeme II



Quelle: Typke, F. Wiering, R. C. Velkamp: A survey of music information retrieval systems. (2004)

Fazit

- Die Inhaltsbasierte Musiksuche basiert auf der Analyse des Musik-Contents
- Anfragentypen:
 - direkte Angabe von Musik-Features
 - *Query by humming*
 - *Query by example*
 - *Query by beatboxing*
- Probleme
 - Meistens Beschränkung auf die Suche der monophonischen Musik
 - Autorenrechte! Musik darf nur legal in den Datenbanken gespeichert werden.

Literatur

- L. Lu, H You, HJ Zhang: A new approach to query by humming in music retrieval (2001)
- Typke, F. Wiering, R. C. Veltkamp: A survey of music information retrieval systems (2004)
- A. Kapur, M. Benning, G. Tzanetakis: Query-by-beat-boxing: Music retrieval for the DJ (2004)
- Deutsches Forschungszentrum für Künstliche Intelligenz:
http://www.dfki.uni-kl.de/KM/content/e13/e109/index_ger.html
- H. Harb, L. Chen: A query by example music retrieval algorithm (2003)
- B. S. Manjunath, P. Salembier, T. Sikora: Introduction to MPEG-7: Multimedia Content Description Interface (2002)