

## Symposium Musikinformatik

Michael Clausen, Frank Kurth, Meinard Müller  
{clausen, frank, meinard}@cs.uni-bonn.de

Die Bereiche *Musik* und *Audio* haben sich in den letzten Jahren zu wichtigen Anwendungsdomänen der Informatik entwickelt. War der informatische Forschungsschwerpunkt bei audiovisuellen Datentypen bisher eher in den Bereichen Bild und Video angesiedelt, so sorgten vor allem Entwicklungen in den Bereichen Audiokompression und Music Information Retrieval (MIR) für einen verstärkten Einsatz der Informatik im Musikbereich. Hier sind neben erfolgreichen Entwicklungen im Standardisierungsbereich (z.B. MPEG-1 bis -7) auch neue Technologien für die Musiksuche und -identifikation zu nennen. Neben den Themenkomplexen Retrieval und Kompression haben sich unter anderem die Gebiete Musik- und Audio-Content-Management, Musikinformationssysteme und -bibliotheken, Automatische (symbolische) Musikanalyse, Audio Pattern Recognition, Annotation und Metadaten, Musikkategorisierung und Recommendation gebildet. Das Symposium *Musik-informatik* soll dazu dienen, deutsche und europäische Forschungsaktivitäten auf diesem noch relativ jungen Gebiet der Informatik näher zusammenzubringen und Synergiepotenzial der zum Teil sehr unterschiedlichen Forschungsansätze herauszuarbeiten. Aus diesem Grund besteht das Symposium aus einer halbtägigen Sitzung mit Vorträgen eingeladener Spezialisten, einer halbtägigen Sitzung mit eher informellen Kurzpräsentationen zu aktuellen Forschungsprojekten, sowie einem abschließenden Rundgespräch.

Bei den eingeladenen Vorträgen bildet die Retrievalkomponente einen deutlichen Schwerpunkt. Im Vortrag *Music Information Retrieval mit Transportdistanzen: Abstandsmaß, Indexierung und Vergleich der Ergebnisse mit menschlichen Vorstellungen von melodischer Ähnlichkeit* (Tytko) wird die Earth Mover's Distance (EMD) als Abstandsmaß für Melodien vorgestellt. Die EMD ist hier unter anderem vielversprechend, weil sie beliebige Kombinationen aus monophoner und polyphoner Musik zulässt und robust gegen Schwankungen in der Tonhöhe und im Tempo, wie sie typischerweise bei "Query-by-Humming" auftreten, ist. Im Vortrag wird gezeigt, wie die EMD für notierte Musik eingesetzt werden kann, wie gut sie mit der Vorstellung menschlicher Experten von melodischer Ähnlichkeit übereinstimmt, und wie ein Index die Suche in der Datenbank beschleunigen kann. Der Vortrag *Inhaltsbasierte Musiksuche in einer multimodalen Systemarchitektur* (Schuller, Rigoll) untersucht als Ergänzung zu traditionellen Suchtechniken (z.B. listenbasierte Selektion und Schlüsselwortsuche) verschiedene inhaltsbasierte Anfragetechniken im Musikretrieval, unter anderem die Suche durch Vorsingen und Vorspielen. Der Vortrag konzentriert sich auf Ansätze zur Realisierung solcher intuitiver Datenbankabfragen unter Verwendung geeigneter Verfahren des maschinellen Lernens. Darüber hinaus werden Probleme und Grenzen bei der Einbindung dieser Verfahren in ein reales Systemumfeld, wie die kontinuierliche Diskrimination von Sprache, Gesang und Musik aufgezeigt. Den Rahmen bildet dabei ein multimodales Interface, das dem Benutzer eine Ausschöpfung der genann-

ten Suchformen durch akustische und manuelle Interaktion ermöglicht.

Eine wichtige allgemeine Zielsetzung von Arbeiten im Bereich Musikretrieval und Musikinformatik im weiteren Sinne ist die automatische Erzeugung von Metadaten, die zur Charakterisierung der einzelnen Musikstücke verwendet werden können. Solche Metadaten unterstützen das Ordnen, Abspeichern, Suchen und Vergleichen in großen Datenbanken. Im Vortrag über *Merkmalsätze für die automatische Klassifikation von Musik- und Audiosignalen* (Kohlrausch, McKinney, Jeroen Breebaart) werden einige Ergebnisse zu Arbeiten zur automatischen Klassifikation von Musiktiteln vorgestellt. Dabei konzentriert sich die Analyse auf den Vergleich verschiedener low-level Merkmale sowie zweier psychoakustischer Merkmalsätze, die aus den Musiksignalen berechnet werden.

Neben der technischen Seite des Musikretrievals und der Merkmalsextraktion geht es in zwei weiteren Vorträgen um Systeme zur Musik- und Audioverarbeitung. Der Vortrag *The Power of Users and Context in Music Information Retrieval* (Baumann) konzentriert sich dabei auf den Retrieval-Bereich und untersucht die soziokulturelle Kompatibilität von MIR Systemen. Hierzu werden eine Systemarchitektur und Beispielapplikationen beschrieben, die im Feldversuch für mobiles Musikretrieval eingesetzt wurden. Anwender wurden hierbei nicht unter den sonst üblichen Laborbedingungen, sondern in ihrem konkreten Kontext beobachtet. Ein weiteres wichtiges Thema im Umfeld MIR stellen kollaborative Ansätze zur Metadaten- und Ontologiegenerierung durch Anwender dar. Der Vortrag beleuchtet hier aktuelle Open Source Projekte und kommerzielle Anwendungen.

Die Repräsentation musikalischer Informationen in Software und Dateiformaten ist bisher sehr spezialisiert angegangen worden. Es gibt eine Vielzahl von Formaten für sehr verschiedene Anwendungsfälle und auf verschiedenen Repräsentationsebenen, z.B. Klangspeicherung (Wav, MP3), Klangerzeugung (CSound, MPEG4 SAOL), Performance (MIDI, MPEG4 SASL), Notation (Finale, MusicXML, Humdrum), Harmonik (ChordML) und Metadaten in verschiedenen Varianten (MPEG7, RDF/DC, Bibliotheksformate). Alle diese Formate und die entsprechende Software sind jeweils auf eine, höchstens zwei Repräsentationsebenen konzentriert, wodurch die Möglichkeit der Verbindung verschiedener Ebenen fehlt. Die *MUSITECH Umgebung für Musikanwendungen* (Weyde) beruht auf einem Objektmodell, das in Java und dem integrierten XML Datenformat MusiteXML repräsentiert werden kann und bietet die Möglichkeit, Musik in verschiedenen Repräsentationen sowie deren Beziehungen zu speichern.

In den Kurzpräsentationen werden Projekte aus unterschiedlichen Bereichen der Musikinformatik vorgestellt. Dies umfasst ein breites Themenspektrum, das sich von Melodieextraktion durch Quellentrennung (Vembu), Modellierung von Gedächtnisprozessen für die computerbasierte Melodieanalyse (Lüsse), über Score Following und Musiksynchronisation (Kaprykowsky), Klassifikation von Musik-Genres mittels Perkussivität (Skowronek, McKinney), eine Untersuchung zur Wahrnehmung von Phrasengrenzen in Pop-Musik (Bruderer, McKinney, Kohlrausch) bis zur Präsentation des SyncPlayer-Frameworks zum inhaltsbasierten Zugriff auf Musikdaten (Fremerey, Damm) erstreckt.

Michael Clausen, Frank Kurth, Meinard Müller  
(Universität Bonn)