

WIE SICH MELODIEN FINDEN LASSEN

JAN-MARK BATKE

Not macht erfinderisch, dachte sich wohl der Mann, der auf einen Passanten zustürzte, ihn fragte »Entschuldigen Sie, kennen Sie...« und ihm daraufhin Folgendes vorsang: »Schnief, schnief di schneuf, schnippel-dischnief, dischneuf«. Gesucht ist offensichtlich der Titel eines Liedes, von dem weder Interpret noch Komponist, sondern nur die Melodie bekannt ist. Diese Szene, die in einer Folge aus Nonstop-Nonsense mit Dieter Hallervorden zu sehen war, ist längst Fernsehgeschichte. Heute ist es näher liegend, eine Suchmaschine zu befragen.

Wer aber eine Suchmaschine wie Google nutzt, trifft auf ein Problem: um den unbekannten Titel zu beschreiben, ist lediglich die Eingabe von Text möglich; andere Eingabevarianten bieten Google & Co. nicht an. Sind Titel und Interpret wie im obigen Szenario aber unbekannt, wird es schwierig, eine Melodie nur in Textform zu beschreiben. Einfacher wäre es, eine Melodie zu singen oder zu summen.

Suchen durch Summen

Dieses Problem lösen spezialisierte Musiksuchmaschinen, die so genannten »Query by Humming«-Systeme (QBH) – was übersetzt »Anfragen durch Summen« bedeutet: Der Anwender summt die gesuchte Melodie in ein Mikrofon, das an einen Computer angeschlossen ist. Software zeichnet das Signal auf, wertet es aus und führt mit dem Ergebnis der Analyse eine Datenbankrecherche durch. Das Ergebnis dieser Recherche wird schließlich als Trefferliste präsentiert.

Für den Erfolg einer Melodiesuche ist es von entscheidender Bedeutung, wie Suchanfrage und Datenbankbestand miteinander verglichen werden. Bei den meisten dieser Melodiesuchmaschinen wird eine symbolische Melodiedarstellung verwendet, welche den gespeicherten Musiktitel repräsentiert. Die Suchanfrage des Nutzers wird vom QBH-System in eine ebensolche symbolische Melodiedarstellung umgewandelt und dann mit dem Datenbankbestand verglichen. Das Ergebnis der Suchanfrage ist eine Liste von Titeln, denen die ähnlichsten symbolischen Melodie-

darstellungen zugeordnet sind. Das Musikstück zur Melodie selbst kann über Titel und Interpret, Noten oder als Aufnahme in Form einer Audiodatei in der Datenbank repräsentiert werden. Audiodateien sind die für Musikdatenbanken das am meisten verwendete Medium und in großer Anzahl als Wave- oder MP3-Datei im Internet vorhanden. So stellt zum Beispiel allein das Internetportal mp3.de [2] über 120.000 frei erhältliche Titel zur Verfügung. Beim kommerziellen Anbieter Musicline werden online über 1,8 Millionen Titel [3], über das Internetportal iTunes mehr als 1 Millionen Titel [4] angeboten.

Musiksuchmaschinen

Die QBH-Systeme gehören zu einer ganzen Gruppe von Systemen, die es ermöglichen nach Musik zu suchen. Diese Music-Information-Retrieval (MIR)-Systeme [1] lassen sich nach den Anfrageformen unterscheiden, die sie dem Nutzer bieten, um die gesuchte Melodie, Musik oder das Lied zu charakterisieren.

- Durch *Text* lassen sich Informationen über Titel, Interpret oder Genre eines gesuchten Musikstücks beschreiben. So suchen herkömmliche Suchmaschinen wie Google, aber auch die meisten Musikportale.
- Query by Humming (QBH) ermöglicht die Suche nach einer Melodie durch *Summen*. Der Nutzer summt die gesuchte Melodie auf Silben wie »na« oder »da«, die über ein Mikrofon aufgezeichnet werden. Für einige Systeme wird auch *Pfeifen* vorgeschlagen, dann wird von Query by Wistling (QBW) gesprochen. Spezieller ist Query by Tapping (QBT), bei dem der Nutzer den Rhythmus des gesuchten Stückes durch Händeklatschen oder *Klopfen* vorgibt.
- Query by Example (QBE) bietet die Möglichkeit, dem Suchsystem eine *Stichprobe* eines vorhandenen Musikstücks darzubieten. Diese Probe kann zum Beispiel ein kurzer Ausschnitt eines Musiktitels sein, der mit dem Mobiltelefon in der Diskothek aufgenommen wurde. Bei der Verarbeitung der Probe wird ein »Audio-Fingerabdruck« ermittelt; man spricht deshalb auch von »Audio-Fingerprinting«.
- Da Musik über *Notenschrift* dargestellt werden kann, ist die Eingabe von Noten selbstverständlich ein geeignetes Mittel zur Melodiesuche. Allerdings setzt sie musikalische Grundkenntnisse voraus und bietet sich daher nur für eine kleine Nutzergruppe an.

Alle genannten Anfragemöglichkeiten lassen sich miteinander kombinieren; so ist beispielsweise ein gesummter Vortrag mit dem Textthinweis auf das Genre Popmusik geeignet, um die Trefferliste einer Datenbankabfrage einzugrenzen.

Musicline und Musipedia

Musiksuchmaschinen unterscheiden sich durch die angebotenen Methoden zur Suche und der internen Repräsentation des Datenbankbestandes; zwei Musikportale mit unterschiedlichen Suchsystemen sind Musicline und Musipedia.

Das Portal *Musicline* [3] ist ein kommerzielles Angebot der Phononet GmbH, das QBH-System wurde vom Fraunhofer-Institut IDMT entwickelt. Es bietet ein Java-Applet an, über das die Nutzer eine gesungene Suchanfrage stellen können. Andere Anfrageformen sind nicht möglich, wodurch es sich um ein reines QBH-System handelt. Die Datenbank enthält vollständige Melodiebeschreibungen und -konturen, die gemäß des MPEG-7-Standards (siehe unten) gespeichert werden können. Das QBH-System von Musicline bietet etwa 3.500 Titel an, hauptsächlich aus dem Genre Pop-Musik.

Musipedia [5] bezeichnet sich als offene Musik-Enzyklopädie und ist angelehnt an das Projekt Wikipedia [6]. Wie dort können Nutzer die Datenbankeinträge editieren und erweitern. Speziell für Melodien wird der im Folgenden erläuterte Parsons-Code verwendet, der sich zur Suche auch direkt als Text eingeben lässt. Alternativ können die Nutzer ein Java-Plugin benutzen, um eine akustische Eingabe in den Parsons-Code zu übersetzen. Sie können zudem das Genre des gesuchten Titels spezifizieren und damit die Ergebnisliste eingrenzen. Die Datenbank enthält etwa 10.000 klassische Titel, 1.000 Popmusikstücke und 15.000 Volkslieder. Die beiden Suchmaschinen unterscheiden sich also durch ihre Nutzerschnittstellen und ihre interne Melodierepräsentation.

Musik digital

Die Art und Weise der Suche nach Musik im Internet hängt sehr stark davon ab, in welcher Gestalt die Musik dargestellt sein soll. Meist werden Audio-Dateien gesucht, aber mitunter auch Noten zu einer Melodie. Die Darstellung von Musik kann folglich in zwei Kategorien unterteilt werden: In der *Audio-Repräsentation* werden Aufnahmen von Konzerten oder Studioproduktion festgehalten, die zum Beispiel als WAV- oder MP3-Datei auf einem Rechner digital gespeichert werden. Die *symbolische Repräsentation* von Musik hingegen meint die klassische Noten-

schrift oder technische Formate für die Musiknotation wie Niff (Notation Interchange File Format) oder Guido (ein generisches Format). Über den Musical-Instrument-Digital-Interface-Standard (MIDI) können Informationen zu Noten, wie sie auf einem elektronischen Musikinstrument gespielt werden, in digitaler Form festgehalten werden.

Um Melodien für QBH-Systeme zu beschreiben, wird häufig eine symbolische Musikrepräsentation gewählt. Dies kann jedoch problematisch sein: Für einen befragten Passanten oder den Bar-Pianisten ist es kein Problem, wenn die gesuchte Melodie in einer anderen Tonart oder in einem anderen Tempo, eventuell mit einigen Lücken oder Ergänzungen, vorgetragen wird. Für eine Suchmaschine wird unter diesen Bedingungen der Vergleich aber schwierig. Die genaue Beschreibung einer gesumten Melodie durch Noten kann daher zur Folge haben, dass eine Suchanfrage erfolglos bleibt, weil die Noten der Suchanfrage und die der Melodie in der Datenbank nicht übereinstimmen oder einander nicht zugeordnet werden können.

Eine Möglichkeit dieses Problem zu umgehen, ist die Wahl einer unschärferen, weniger eindeutigen Melodiedarstellung. Für QBH-Systeme können zum Beispiel Melodiekonturen verwendet werden, die in Form des Parsons-Codes oder als MPEG-7 MelodyContour dargestellt werden.

Parsons-Code

Eine Melodiekontur beschreibt lediglich, ob sich die Tonhöhe einer Melodie nach oben oder unten ändert oder gleich bleibt. Damit sind mindestens drei Symbole notwendig, mit der eine Kontur beschrieben werden muss: »U« (Up), »D« (Down) und »R« (Repeat). So ist eine Beschreibung der Melodiekontur als Zeichenkette möglich. Beispielsweise repräsentiert »*DRUDRDUUUURR« den Anfang des Lieds »Hänschen klein« (siehe Abb.); der Stern (*) steht dabei für die erste Note, der noch kein Konturwert zugewiesen werden kann. Über diese Buchstabenfolgen werden für die Suche nach Ähnlichkeiten zwischen Melodien Verfahren zur Zeichenkettensuche verwendbar, die aus der Informatik bekannt sind. Der Parsons-Code einer Melodie wird dabei als Suchwort aufgefasst und in der Datenbank wird das gleiche oder ein möglichst ähnliches Wort gesucht.

MPEG-7

Der Multimedia-Standard MPEG-7 definiert Methoden zur Beschreibung von Audio- und Videodaten. Er kümmert sich im Gegensatz zu vorangegangenen Standards wie MPEG-1, -2 und -4 nicht mehr um den Inhalt von Audio- und Videomaterial selbst, sondern enthält die »Bits about the Bits« (Thom et al. 1999). Für Anwendungen wie QBH-Systeme

steht das MPEG-7 Melody Description Scheme zur Verfügung (vgl. ISO 2001). Die Beschreibung der darin enthaltenen MelodyContour erfolgt nuancierter als im Parsons-Code über 5 Stufen (-2, 2: eine kleine Terz oder mehr ab- bzw. aufwärts; -1, 1: einen Halb- oder Ganzton ab- bzw. aufwärts; 0: Prime). Darüber hinaus werden im ContourType aber auch Informationen über den Rhythmus festgehalten, indem jeder Melodienote die Nummer des Schlags im Takt zugeordnet wird. Damit können Ähnlichkeitsmaße verwendet werden, die Melodie- sowie Rhythmuskontur vergleichen.



The image shows a musical staff in 4/4 time with a key signature of one flat (B-flat). The melody consists of 12 notes: D4, E4, F4, G4, A4, Bb4, A4, G4, F4, E4, D4, C4. Above the staff, the notes are grouped into four measures, each labeled with a number 1 through 4. Below the staff, the Parsons code and the MPEG-7 MelodyContour are provided.

Parsons: *	D	R	U	D	R	D	U	U	U	U	R	R
MPEG-7 MelodyContour: *	-2	0	1	-2	0	-1	1	1	1	1	0	0

Abbildung: Die Noten zu »Hänschen klein« und die Beschreibung dieser Melodie im Parsons-Code bzw. als MPEG-7 MelodyContour.

Fazit

Mögliche Anwender von QBH-Systemen sind Internet-Nutzer, die nach einem Musiktitel recherchieren, um ihn zum Beispiel als CD oder MP3-Datei zu erwerben. QBH-Systeme werden bereits im kommerziellen Musikvertrieb wie bei Musicline angeboten. Darüber hinaus ist aber auch die professionelle Verwendung von QBH-Systemen denkbar, die zum Beispiel für Musikwissenschaftler zur Analyse großer Musikbestände nützlich sein könnte. Juristen haben bei der Verfolgung von Urheberrechtsverstößen ebenfalls ein Interesse an dem Überblick über einen großen Musikbestand. Je nach Bedürfnissen unterteilt sich die Gruppe der Anwender in Laien und professionelle Nutzer.

Ob QBH-Systeme ein breites Publikum finden werden, bleibt abzuwarten. Derzeit ist die Integration von Java-Applets in die Umgebung des Rechnerbetriebssystems noch nicht so nahtlos, dass eine Audioaufnahme immer mühelos gelingt. Die Erfolgsquote solcher Systeme ist begrenzt, da eine zielführende Suche durch eine verwendbare Suchanfrage bedingt ist. Einfacher gesprochen: Nicht jeder kann so gut singen, wie es für ein QBH-System notwendig ist. Einige Nutzer neigen sogar dazu, eher den Liedtext zu sprechen. Deshalb sollten Musiksuchsysteme der Zukunft vielfältige Eingabemöglichkeiten anbieten, eine sinnvolle Ergänzung wären sicherlich auch Bilder zu CD-Covern oder Filmen.

Musicline verzeichnet einige Hundert Anfragen pro Tag, längst nicht immer sind die gesuchten Melodien auch in der Datenbank enthalten. Einen möglichen Weg bestehende Datenbankbestände zu vergrößern, zeichnet Musipedia vor. Dort haben Nutzer die Möglichkeit, fehlende Melodien zu ergänzen und machen davon auch Gebrauch. Festzuhalten ist, dass der klassische Handel mit Tonträgern sich in Richtung Internet-handel mit Dateien verschiebt. Apple gibt für sein Portal iTunes 35.000 verkaufte Titel pro Tag an, die Konkurrenz, u.a. auch Microsoft, folgt mit eigenen Angeboten. Unter diesem Gesichtspunkt kommt MIR- und speziell QBH-Systemen eine große Bedeutung zu.

Literatur

International Organization for Standardization (ISO) (2001): *Information Technology Multimedia Content Description Interface Part 4: Audio*. 15938-4.

Thom, David/Purnhagen, Heiko/Pfeiffer, Silvia und MPEG Audio Subgroup (1999): »MPEG Audio FAQ«. [WWW-Dokument www.tnt.uni-hannover.de/project/mpeg/audio/faq - Zugriff 31.10.2004].

Digitale Verweise

[@1] <http://mirsystems.info>

[@2] www.mp3.de

[@3] www.musicline.de

[@4] www.itunes.com

[@5] www.musipedia.org

[@6] www.wikipedia.org