

Organisation von privaten digitalen Musikbibliotheken

Philip Koene
philip.koene@ifi.lmu.de

Universität München
Medieninformatik
Amalienstrasse 17, 80333 München, Deutschland

Zusammenfassung Durch die zahlreichen Möglichkeiten der Beschaffung von digitaler Musik und die Notwendigkeit digitale Musik zu besitzen, um sie auf einem portablen MP3-Player abspielen zu können, ist die Anzahl an privaten digitalen Musikbibliotheken in starkem Wachstum begriffen. Die unterschiedlichen Quellen von digitaler Musik führen aber auch zu sehr unübersichtlichen Sammlungen. Es stellt sich die Frage, wie man dem Musiksammler das Organisieren und Browsen (dt. Durchblättern) seiner Bibliothek erleichtern kann. Zur Beantwortung dieser Frage gibt diese Arbeit einen Überblick über die Methoden der Metadaten-Akquise und Verifikation und beschreibt, wie man diese zur Organisation von Musiksammlungen nutzen kann. Außerdem greift sie einige Nutzerbefragungen zu diesem Thema auf und umreißt abschließend die Möglichkeiten der Visualisierung von Musikbibliotheken.

1 Einleitung

Mit der Markteinführung des iPods von Apple am 23.10.2001 traten die portablen MP3-Player ihren Siegeszug an [22]. Der iPod erreichte alleine im Jahr 2005 mit seinen diversen Produktlinien Verkaufszahlen im zweistelligen Millionenbereich [19]. Neben Apple konkurrieren unter anderem Creative, Rio, Samsung, Sony und Reigoncom [20] in diesem nahezu exponentiell anwachsenden Markt, welcher sich Prognosen zufolge bis zum Jahr 2009 vervierfachen soll [18]. Die einzige Möglichkeit Musikdaten auf all diese tragbaren MP3-Player zu überspielen, funktioniert via Heimcomputer. Die Musik für den portablen Hörgenuss muss also zunächst in digitaler Form auf der Computerfestplatte des Nutzers als Musiksammlung vorhanden sein.

Selbst wenn man alle anderen Beweggründe außer Acht lässt, welche einen Nutzer zur Erstellung einer persönlichen digitalen Musikbibliothek motivieren können, so bescheinigen doch alleine diese Zahlen, dass die Anzahl der Musiksammlungen auf den privaten Heimcomputern bereits jetzt gewaltige Ausmaße angenommen hat und in Zukunft weiter wachsen wird.

Die Quellen, aus welchen der Nutzer heutzutage beim Aufbau einer solchen Musikbibliothek schöpfen kann, sind vielfältig. Sie reichen vom Digitalisieren gekaufter Musik-CDs über das Austauschen von Musikstücken mit Freunden und das größtenteils illegale Austauschen von MP3s über das Internet mit "Freunden" aus aller Welt, bis hin zu den mittlerweile immer weiter verbreiteten Online-Musikläden, deren Marktführer Apple mit iTunes gegen Ende des Jahres 2005 eine Milliarde Songs verkauft hat [21].

So unterschiedlich die Quellen digitaler Musik sind, so unterschiedlich sind leider oft auch die Normen, welche bei der Benennung der einzelnen Songdateien zur Anwendung kommen. Bei einer Bibliotheksgröße von wenigen hundert Musiktiteln mag trotzdem eine Zugriffsmethode über Dateilisten und Ordnern noch ausreichend schnell und zielsicher für den Nutzer sein, sie beginnt jedoch zu versagen, wenn man auf tausende oder sogar zehntausende Musikstücke trifft. Die meisten Bibliotheksprogramme bieten dem Nutzer, um sich in solch einer großen Sammlung zurecht zu finden, eine Suche mittels Eingabemaske an. Dieses System belässt jedoch die Informationslast bei dem Nutzer. Er muss zumindest den korrekten Interpreten, Songtitel oder einen Albumnamen haben, um zu dem gewünschten Ziel zu gelangen. Darüberhinaus bieten solche Systeme keine weiterführende Information über die gefundene Musik oder einen Kontext wie Genre oder Erscheinungsjahr an.

Es stellt sich also die Frage, wie man der zunehmenden Zahl an Sammlern digitaler Musik das Organisieren und Browsen ihrer Bibliotheken erleichtern kann und welche Zusatzleistungen und -informationen man ihnen bieten möchte.

Zur Beantwortung dieser Frage werde ich in der folgenden Arbeit zunächst einmal im folgenden Abschnitt einen Überblick über verwandte Literatur auf dem Gebiet der Musik Metadata (dt. Metadaten) geben, in Kapitel 3 werde ich dann exemplarisch eine Anwendung namens pTunes vorstellen, welche die meisten Konzepte eines umfangreichen Bibliotheksprogrammes beinhaltet. Kapitel 4 behandelt ein System zur automatischen Erkennung des Genres eines Musikstückes und dessen Umsetzung in ein intuitiv bedienbares Kartenprogramm. Kapitel 5 und 6 befassen sich mit zwei wichtigen Zusatzthemen zu diesem Gebiet, zum einen werden Nutzerbefragungen zu Musiksammlungen analysiert, zum anderen werden neue Möglichkeiten der Visualisierung einer Bibliothek umrissen. Abschließend fasst ein kurzes Fazit die Arbeit noch einmal zusammen.

1.1 Grundlegende Begriffe - MP3 und ID3-Tag

Das MP3-Dateiformat ist heute das mit großem Abstand am weitesten verbreitete Dateiformat für digitale Musik [14] und besticht vor allem durch hohe Kompressionsraten bei geringem Qualitätsverlust. Allerdings hat es, vom Fraunhofer Institut zunächst hauptsächlich zu Ausstrahlungszwecken entwickelt, in seinem Datei-Header nahezu keinen vorgesehenen Platz für Metadata, lediglich 1-Bit Felder für *copyright*, *copy/master* und *private*. Im Jahr 1996 entwickelte Eric Kemp unabhängig davon den ID3v1-Tag [12][13], eine Art Anhänger, welcher Metainformation enthielt und mittels eines Programmes an jede beliebige Datei angehängt werden konnte. Diese Etikettierung fand schnell Zuspruch bei Musiksammlern aus aller Welt als Hilfe zur Organisation ihrer MP3s. Heutzutage ist der ID3-Tag in seiner mittlerweile zweiten Version ID3v2 die gängige Verfahrensweise um Musikmetadata zu speichern und wurde auch in zahlreicher Software im Bereich digitaler Musik integriert. Mit dem Sprung auf die zweite Version wurde unter anderem der ID3-Tag an den Anfang des Musikfiles gestellt, womit schnelles Durchsuchen der Tags möglich war, ohne die Informationen des MP3s selbst lesen zu müssen. Außerdem wurde die Möglichkeit eröffnet unbegrenzt neue Metadatafelder anzulegen [2].

2 Verwandte Arbeiten

Aufgrund der Popularität von digitaler Musik sind natürlich bis zum heutigen Tage schon einige "Jukebox"-Systeme entstanden, welche sowohl das Verwalten von Musikbibliotheken als auch das Abspielen von MP3s beinhalten. Die bekanntesten hiervon sind wohl iTunes, WinAmp und MP3 Voodoo, wobei der Verwaltungsanteil bei den genannten Programmen sehr unterschiedlich ausgeprägt ist. Während WinAmp in erster Linie ein Player für digitale Musik ist und als Verwaltungsfunktion nur die Playlist anbietet, implementieren iTunes und MP3 Voodoo mehrere Ansichten und Sortiermöglichkeiten, etwa nach Interpret oder Album. In akademischer Hinsicht wurde in der digitalen Musik vor allem der Bereich des MIR - Music Information Retrievals im Detail betrachtet, daher gibt es unzählige, größtenteils sehr technisch orientierte Arbeiten zu verschiedenen Algorithmen, welche Query-by-humming möglich machen. So beschreibt etwa Steffen Pauws in [15] die einzelnen Schritte zur Tonlagenerkennung für ein solches Query-by-humming System (dt. Suche durch Vorsummen). Aber auch im Bereich der digitalen Musikbibliotheken gibt es mittlerweile eine nennenswerte Anzahl von Arbeiten; Mahia Parparita und Ben Kernighan haben etwa in [1] den Ansatz gemacht, ein eigenes, möglichst umfassendes Bibliotheksprogramm namens pTunes zu erstellen und James Howison und Abby Goodrum haben in [2] den Versuch unternommen, die Entwicklung der Musikmetadata im Vergleich mit Metadata zu akademischen Arbeiten darzustellen. Ein großer Teil der eher technisch angelegten Arbeiten auf diesem Gebiet befasst sich mit der automatischen Genre-Erkennung von Musikstücken. So greifen etwa Rauber et al. in ihren Arbeiten [5],[6] und [7] ein von Elias Pampalk 2001 entwickeltes Prinzip der "Islands of Music" [16] auf, in dem Musikstücke nach ihrer Ähnlichkeit im Genre auf einer SOM (Self Organizing Map), einer Karte der Musiksammlung platziert werden.

Einmal im Jahr findet in Izmir eine Konferenz zu Music Information Retrieval statt, in der auch immer wieder beachtenswerte Arbeiten zu dem Thema Musikbibliotheken veröffentlicht werden; so wurden etwa auf der Konferenz im Jahr 2004 mehrere Nutzerbefragungen herausgegeben [10][11] und Marc Torrens et al. präsentierten in [4] mehrere neue Konzepte zur 2D-Visualisierung von Musiksammlungen.

3 Konzepte eines Bibliothekprogramms - pTunes

Bei der Entwicklung eines Bibliothekprogrammes für digitale Musik gibt es einige problematische Aspekte, die es von den Designern zu lösen gilt, aber auch manche Konzepte, deren Implementierung aus dem resultierenden Programm mehr macht als eine Suchmaske, deren Felder vom Nutzer mit Informationen gefüllt werden müssen. Mihai Parparite und Brian Kernighan haben bei der Entwicklung ihres Programmes pTunes, vorgestellt in [1], einige dieser Probleme, zum Beispiel die Herausforderung der Datenuniformität gelöst, und durch die Integration von Online-Diensten wie CDDB Mehrwert für den Nutzer geschaffen. Getestet haben die Entwickler ihr Produkt im April des Jahres 2003 anhand der Musik, welche im Netzwerk der Princetown University zu finden war. 86.072

Musikdateien, von 209 unabhängigen Nutzern zusammengetragen, erlaubten es den Autoren die Effektivität ihres Produktes bei der Integration mehrerer unterschiedlicher Musiksammlungen in eine homogene Musikbibliothek zu zeigen.

3.1 Normalisierung von Dateinamen und Metadata

Um die Möglichkeit zu bieten den Inhalt einer Musikbibliothek nach Songname, Interpret, Album etc. zu ordnen, beziehungsweise diese Metadaten als Suchkategorien verfügbar zu machen, muss das Bibliotheksprogramm natürlich erst einmal an diese Information - so vorhanden - in einer gegebenen Musiksammlung gelangen. Leider kann jedoch die Art und Weise, in welcher diese Informationen gespeichert sind, sehr unterschiedlich sein. Die meisten MP3s haben glücklicherweise einen angehängten ID3-Tag, in dem Metadata wie Interpret, Album, Titel, Genre, Entstehungsjahr usw. gespeichert sind. Bei vielen Musiksammlungen sind diese ID3-Tags jedoch entweder nicht vorhanden, oder unausgefüllt, da der Sammler sich dazu entschlossen hat, die Information auf anderen Wegen zu speichern. Eine weit verbreitete Methode ist etwa das Einordnen der einzelnen MP3s in eine hierarchische Ordnerstruktur, die zum Beispiel nach Interpret und Album benannt und gegliedert ist. Auch in den Namen der Musikdateien selber ist nahezu immer Metainformation gespeichert. Glücklicherweise finden sich hierbei sehr oft bestimmte Schemata, nach denen sich die Dateinamen zusammensetzen, wie z.B. "<Interpret> - <Titel>.mp3" oder "<Track Nummer>. <Interpret> - <Album> - <Titel>.mp3". Diese Schemata stammen oft von der Software, welche ursprünglich verwendet wurde um das MP3 zu kodieren, und die, vom Nutzer eingegebene, Metadata so umgesetzt hat. Die erste Aufgabe ist es also zunächst, aus der vorhandenen Musiksammlung diese Information zu extrahieren und in geeigneter Form zu speichern. pTunes macht dies, indem es die Dateien nach "artist", "album" und "titel" indiziert und diese Indizes separat speichert. Alternativ könnten diese Informationen auch in neu angelegten Feldern des ID3-Tags gespeichert werden, was den Vorteil hat, dass sie sich mit der Musikdatei mitbewegen. Um redundante Stördaten für den Nutzer zu vermeiden, muss ein Weg gefunden werden, doppelte Einträge herauszufiltern. Diese entstehen jedoch meistens durch Variationen, ausgelöst durch Rechtschreibfehler, falsche Buchstabierung oder Abkürzungen. pTunes normalisiert zu diesem Zweck zunächst einmal die gefundenen Informationen, das heißt, alles wird in Kleinbuchstaben umgewandelt und Leerstellen, Vorsilben sowie Nachsilben werden entfernt. Ähnliche und dadurch möglicherweise doppelte Einträge werden dann mittels eines in Perl implementierten Algorithmus herausgefiltert. Da dieses System jedoch nicht wissen kann, dass etwa "Breeder" und "Breeders" zwei verschiedene Bands sind, haben die pTunes-Entwickler eine kleine Applikation geschrieben, anhand derer vom Anwender Doppelmeldungen überprüft werden können.

3.2 Integration von Online-Diensten

Eine Möglichkeit, an schlicht fehlende Metainformation von Musikdateien zu gelangen, ist die Integration von speziellen Online-Diensten in das Bibliotheks-

programm, welche selbige Information zur Verfügung stellen - einen Internetanschluß beim Nutzer natürlich vorausgesetzt. Diese Online-Dienste, z.B. CDDB, FreeDB und MusicBrainz basieren darauf, dass sie jedem, der sie in Anspruch nimmt, die gewünschte Metainformation, sofern vorhanden, gewährleistet. Fehlt die Information jedoch in der Datenbank, so wird der Nutzer gebeten, diese seinerseits preiszugeben, um somit die Datenbank zu erweitern. Es besteht auch die Option, einen Online-Dienst in Anspruch zu nehmen, der seine Datenbank aus kommerziellen Zwecken führt, wie etwa Amazon.com - eine Website, die Metainformation für die Käufer von Musik-CDs im eigenen Online-Shop unterhält. Diese Dienste haben den Vorteil, aus kommerziellen Gründen recht akkurat zu sein und auch weiterführende Informationen wie etwa Album-Covers oder Kundenrezensionen anzubieten, sie sind aber unter Umständen kostenpflichtig.

CDDB und FreeDB: Das System CDDB wurde im Jahr 1996 entwickelt und trägt heute den Namen Gracenote. Im Einklang mit der 1996 vorhandenen Welt der digitalen Musik ist CDDB ganz auf die CD als primäres Trägermedium ausgelegt. Es identifiziert einzelne Musikdateien anhand einer Prüfsumme ihrer Dateilänge. Diese muss nicht notwendigerweise einzigartig sein, unter Berücksichtigung der Reihenfolge einzelner Musikdateien auf einer gekauften Original-CD aber wird die Identifizierung sehr akkurat. CDDB bietet dann entweder die verfügbare Metainformation an und bittet den Nutzer diese zu verifizieren, oder der User wird aufgefordert, fehlende Information zur Datenbank hinzuzufügen. Dieses System hat natürlich den Nachteil, dass es ein Musikstück nur identifiziert, wenn es auf die Original-CD zugreifen kann, also üblicherweise bei dem "Rip"-Vorgang, dem Enkodieren einer CD. Aus dem Internet heruntergeladene oder anderweitig erworbene MP3s kann es nicht erkennen. Unglücklicherweise haben die Besitzer von CDDB bei der Umbenennung des Services in Gracenote die angesammelte Album-Metainformation als ihr geistiges Eigentum deklariert und begonnen den Kunden einen Geldbetrag für die Nutzung abzuverlangen. Allerdings wurde als Antwort auf diese Kommerzialisierung mit FreeDB eine Open-Source Version des Dienstes ins Leben gerufen.

MusicBrainz: Der Online-Dienst MusicBrainz ist in seiner Konzeption ähnlich wie CDDB, benutzt jedoch eine Prüfsumme der tatsächlichen musikalischen Eigenschaften einer MP3 als Identifikator, was natürlich eine gewisse Soundanalyse notwendig macht. Der Vorteil dieser Vorgehensweise ist, dass jede beliebige Musikdatei in die Datenbank aufgenommen werden kann, ohne eine Original-CD zu benötigen. Leider ist die Soundanalyse etwas rechenaufwendiger als bei CDDB und auch die Präzision der Identifikation von Songs ist bei weitem nicht so groß wie bei dem CD-basierten Vorgänger. MusicBrainz ist allerdings in seiner Anwendung wesentlich einfacher für den Nutzer als etwa CDDB, da es einfach die ID3-Tags einer Musiksammlung ausliest und dann wechselseitig die fehlenden Informationen durch die in der eigenen Datenbank ergänzt.

Amazon.com: Die Webseite Amazon.com ist, wie bereits erwähnt, primär ein Online-Shop für z.B. Musik-CDs und unterhält zu diesem Zwecke auch eine Datenbank der einzelnen, zum Verkauf stehenden, Musikalben. Amazon bietet Außenstehenden die Möglichkeit, über ihren sogenannten "Amazon Web Service"

Anfragen an diese Datenbank zu stellen, welche mittels eines XML-Dokumentes mit allen zur Verfügung stehenden Metadaten über ein Produkt, inklusive Albumcover und Kundenrezensionen beantwortet werden.

3.3 Methoden der Genre-Analyse

Das Genre eines Musikstückes ist ein wichtiger Anhaltspunkt für Nutzer, die zwar eine bestimmte Art von Musik aus ihrer Sammlung hören wollen, jedoch im Moment keinen Songtitel, Albumnamen oder Interpreten im Kopf haben. Darüber hinaus sind die Nutzer eines Bibliotheksprogrammes vorraussichtlich mit der Organisation eines Musikladens nach Genre vertraut und finden sich deshalb in diesen Kategorien intuitiv zurecht. Die zuvor genannten Online-Dienste versorgen den Nutzer mit Genre-Informationen und auch im ID3-Tag ist ein Feld für das Genre vorgesehen. Kann eine MP3 jedoch von einem Onlinedienst nicht identifiziert werden, oder ist in dessen Datenbank schlicht nicht vorhanden und der ID3-Tag ist auch noch unvollständig, so bleibt als letzte Möglichkeit der Genreexpression die Sound-Analyse des Musikstückes. pTunes verwendet hierfür ein System namens MARSYAS, welches, wie beschrieben in [17], versucht, die Eigenschaften eines Musikstückes, also Klangfarbe, Instrumentierung und Struktur, zu extrahieren und dann mit 10 Genre-Schablonen zu vergleichen. Es erreicht dabei eine Treffergenauigkeit von sechzig bis siebzig Prozent. Andere Systeme, wie etwa die in Kapitel 4 beschriebene Rhythm-Pattern-Extraktion können deutlich mehr erreichen. Die Entwickler von pTunes sahen sich nach ihren vielfältigen Methoden der Genre-Extraktion jedoch mit 7.780 Genres für 9.994 Interpreten konfrontiert und waren dementsprechend gezwungen eine Art hierarchische Liste von 140 "kanonischen" Genres zu erstellen, welche in Abbildung 1 auschnittsweise zu sehen sind. Die Zuordnung auf die kanonischen Genres nahmen sie dann von Hand vor, was durchaus einige Zeit in Anspruch genommen haben wird. Um die Möglichkeit zu haben, Fehlinformation auszusieben und trotzdem einige Interpreten und Lieder in mehrere Genres einordnen zu können, haben die pTunes-Designer Gewichtungen für Genreeinträge nach Häufigkeit des Eintrages und Glaubwürdigkeit der Quelle vergeben und einen Gewichtungs-Schwellenwert für die Aufnahmen in die Musiksammlung erstellt.

3.4 Interface von pTunes

Bei der Gestaltung der Benutzeroberfläche geht pTunes einen eher konservativen Weg. Es greift die dreiteilige Gliederung von iTunes auf und erweitert sie um eine vierte Ebene, dem Genre. Dementsprechend zeigt die obere Hälfte des User-interfaces, wie in Abbildung 1 zu sehen, drei alphabetisch geordnete Fenster für "Genre", "Artist" und "Album". Die vorhandenen MP3s einer dort selektierten Kategorie werden dann in dem Abschnitt "Title" unten rechts gezeigt, Ergebnisse einer Suchanfrage in der Eingabemaske oben rechts werden auch dort dargestellt. Das Fenster unten links ist für die Anzeige von zusätzlicher, hauptsächlich von Amazon.com stammender Metainformation wie etwa das Albumcover. Das ganze System ist über eine HTML-Schnittstelle für einen Webbrowser ausgelegt, um einerseits eine Kompatibilität mit vielen Betriebssystemen zu erreichen und andererseits das stark verwendete Prinzip der Verlinkung zu unterstützen.



Abbildung 1. Interface von pTunes.

Wie zu sehen, stützt pTunes seine Benutzeroberfläche eher auf konservative und dem Benutzer durch die Verwendung anderer Programme und Betriebssysteme vertraute Konzepte. Die Funktionsweise der GUI ist stark an die von HTML bekannte Verlinkungsstruktur, sowie an Explorer-artige Datei- und Ordnerlisten angelehnt. Es gibt jedoch derzeit auch Entwicklungen, die auf neue Arten der Visualisierung einer Musiksammlung abzielen, wie etwa die Anordnung der Musikstücke auf einer zwei- oder dreidimensionalen Landkarte, in der topographische Unterschiede verschiedene Musikgenres darstellen. Marc Torrens et al. zum Beispiel ordnen, wie in Kapitel 6 beschrieben, MP3s auf einer runden, zoombaren Scheibe, die in Segmente unterteilt ist und erreichen damit eine sehr intuitive Bedienbarkeit und Playlist-Erstellung.

3.5 Weiterführende Servicefunktionen

Die in den vorhergehenden Unterkapiteln beschriebenen Konzepte befassen sich, so wie die Anwendung pTunes, größtenteils mit der Extrahierung, Analysierung und Verifizierung der grundlegenden Metainformationen. pTunes sieht es also als seine grundlegende Aufgabe an, dem Nutzer die Organisation seiner Musiksammlung zu erleichtern. Darüberhinaus versucht es jedoch mit der Einbindung von Information aus der Amazon.com-Datenbank, dem Nutzer einen Mehrwert zu bieten, der nicht essentiell für ihn ist, sondern als Servicefunktion zu verstehen ist. Eine Diskographie zu einem bestimmten Interpreten zu sehen, hilft dem Nutzer eines Bibliotheksprogrammes eher weniger bei dem Auffinden eines

Musikstückes, kann aber von ihm durchaus positiv aufgenommen werden und die Interaktion des Sammlers mit seiner Sammlung vertiefen. In diesem Sinne ist eine Ausweitung dieser weiterführenden Metainformation durchaus denkbar. pTunes enthält schon aus der Amazon.com Datenbank die Empfehlungen von Kunden, welche zu bestimmten Alben und Interpreten auf ähnliche und ähnlich gute Musikstücke verweisen. Die Integration ganzer Kundenrezensionen oder auch einer Biographie und Photos eines Interpreten wäre durchaus denkbar. Die wohl am häufigsten von den Nutzern verlangte und zum Beispiel in iTunes auch schon implementierte Zusatzfunktion wäre die Möglichkeit, Musiktexte einzusehen, sei es um ein nicht verstandenes Textsegment zu überprüfen, oder einfach nur mitzusingen.

Unterstützung von "Thrashing und Sickness": Eine weitere, völlig andere Art der Servicefunktion ist die Sensibilisierung eines Bibliotheksprogrammes auf die Hörgewohnheiten des Nutzers. Sally Jo Cunningham, Matt Jones und Steve Jones [3] zeigen in ihrer Analyse vieler realer CD-Sammlungen zum Beispiel, dass ein relativ kleiner Stapel an CDs sich meistens in der Nähe des Abspielgerätes befindet, während der Großteil der Sammlung archiviert in irgendeinem Regal verweilt. Neue CDs und einige CDs aus der archivierten Sammlung kommen immer wieder zu diesem kleinen Stapel hinzu und werden von dem Musiksammler sehr oft gehört ("Thrashing"), bis eine gewisse Sättigung einsetzt ("Sickness") und die CD zum Ende des Stapels und letztendlich wieder in das Archiv wandert. Ein Bibliotheksprogramm könnte diese Art der Benutzung unterstützen indem es zum Beispiel eine sich ständig aktualisierende, automatisch erstellte Playliste anbietet, die auf das Hörverhalten des Nutzers abgestimmt ist. iTunes macht mit seiner "Smart Playlist"-Funktion schon einen Schritt in diese Richtung.

Unterstützung idiosynkratischer Genres: Häufig wird Musik von den Hörern anhand des Anlasses charakterisiert, bei dem er sie hören will. Dies kann zum Beispiel beim Sport oder der Arbeit sein; manchmal wird aber auch eine Musikauswahl getroffen, um eine bestimmte Stimmungslage zu verstärken oder zu dämpfen. Diese idiosynkratischen Charakterisierungen können sehr unterschiedlich ausfallen und sind deswegen kaum vorherzusehen. Eine Unterstützung dieser Hörweise in einem Bibliotheksprogramm wäre zum Beispiel zu erreichen, indem man es dem Nutzer ermöglicht eigenständig Genres anzulegen und ihnen einzelne Musikstücke, Interpreten oder Alben zuzuordnen. Eine andere, und für den Musikhörer wesentlich unkompliziertere Möglichkeit wäre die Implementation einer "Query-by-example"-Suche, bei der ein Beispielmusikstück gegeben wird und das Programm versucht, mittels Sound-Analyse möglichst ähnliche Songs zur Auswahl zu stellen.

4 Genre-Mapping anhand von Sound-Analyse

Im Vorangehenden wurde an manchen Stellen die Notwendigkeit einer Sound-Analyse angesprochen, um zum Beispiel die anspruchsvolle Problematik der automatischen Genre-Erkennung zu lösen, oder eine "Query-by-example"-Suche

anbieten zu können, die unabhängig von manuell zugewiesener Metainformation operieren kann. Das Kernstück einer jeden Sound-Analyse ist die Extraktion von Audio-Eigenschaften und beeinflusst damit auch nachhaltig die Effizienz der Analyse. Das Ziel der Extrahierung von Audio-Eigenschaften ist es, eine Semantik von Musikstücken zu gewinnen, anhand derer man unterschiedliche Arten von Musik charakterisieren kann. Dabei muss die Datenmenge im Vergleich zu der ursprünglichen Audiodatei drastisch reduziert werden, um Ähnlichkeitsvergleiche in einer anwenderfreundlichen Geschwindigkeit ablaufen zu lassen, aber zugleich sollte genügend prägnante und beschreibende Information übrig bleiben, um diese Vergleiche erst möglich zu machen. Welche speziellen Eigenschaften einer Audiodatei dabei extrahiert werden, ist von System zu System unterschiedlich. Während das zuvor erwähnte MARSYAS System sich auf Klangfarbe und Instrumentierung stützt [17], verwenden die meisten "Query-by-humming"-Umsetzungen eine Melodie-Extraktion. Der im nachfolgenden beschriebene Algorithmus von Rauber et al. [7] verwendet sogenannte "Rhythm Patterns" (dt. Rhythmusstrukturen).

4.1 Rhythm-Patterns

Der Algorithmus für die Extraktion von Rhythm-Patterns ist ein zweistufiger Prozess. Im ersten Teil wird aus der spektralen Repräsentation eines Musikstückes die Lautstärke in bestimmten Frequenzbereichen abgelesen. Dabei werden diverse psycho-akustische Phänomene berücksichtigt, um dem menschlichen Hörapparat Rechnung zu tragen und eine Lautstärke festzustellen, wie sie vom Mensch wahrgenommen wird. Damit wird versucht das menschliche Hörempfinden soweit als möglich nachzustellen, um möglichst authentische Ähnlichkeitsvergleiche anstellen zu können. Im zweiten Schritt werden dann diese Lautstärkenwerte in eine zeitinvariante Form gebracht, aus der sich dann ablesen lässt, wie ausgeprägt die Rhythmen in den einzelnen Modulationsfrequenzen, also bei bestimmten "beats-per-minute"-Werten sind.

Wie in Abbildung 2 beschrieben, wandelt der Algorithmus zunächst einmal die Audiodateien in ein PCM-Mono-Audioformat um; mit einem Input-Signal von 44,1 kHz. Komprimierte Musikdateien wie MP3 werden dabei entpackt. Der Audiotrack wird dann in 6-Sekunden-Segmente zerschnitten, auf denen jeweils eine "Short Time Fast Fourier Transformation" angewendet wird, um ein Frequenzspektrum alle 11,5 ms zu erhalten. Um die Datenmenge zu verkleinern, werden die Frequenzbereiche dieser Spektrogramme auf 24 sogenannte "kritische Bänder" reduziert, welche laut der Bark-Skala für das menschliche Hören von Bedeutung sind. Anschließend wird mit dem "Spektral Masking" der Tatsache Rechnung getragen, dass bestimmte Töne für das menschliche Gehör von vorrausgehenden oder nachfolgenden Tönen auf benachbarten Frequenzbändern verschluckt werden. Daraufhin werden die Lautstärken der einzelnen Frequenzbänder gemäß der menschlichen Lautheitswahrnehmung in Phon und schließlich in Sone umgerechnet.

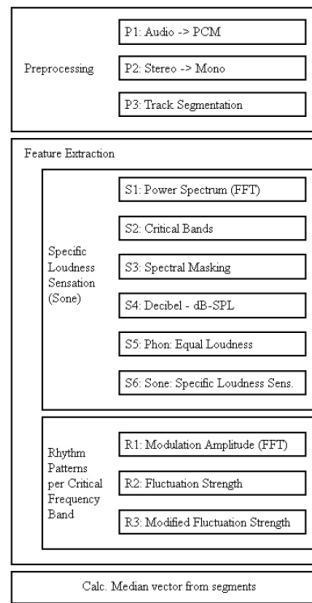


Abbildung 2. Funktionsweise der Rhythm Pattern Extraction.

Um nun aus den 511 Frequenzspektren pro 6-Sekunden-Segment die bereits beschriebene zeitinvariante Repräsentation zu erhalten, wird erneut eine FFT angewendet. Zu guter Letzt wird der Durchschnitt aller 6-Sekunden-Segmente errechnet, um eine Beschreibung des gesamten Musikstückes zu erhalten. Der somit extrahierte Rhythm-Pattern-Deskriptor kann nun zum direkten Vergleich zweier Musikstücke herangezogen werden, und somit auch, den Rhythm-Pattern-Deskriptor eines prototypischen Genrevertreters vorrausgesetzt, zur Genreordnung von Musik verwendet werden. Die Autoren Rauber et al. haben ihren Algorithmus, zusammen mit den in [7] beschriebenen "Rhythm Histogram Features" und dem "Statistical Spectrum Descriptor" an drei Musikdatenbanken, unter anderem in einem Genre-Klassifikations-Wettkampf der Ismir-Konferenz im Jahre 2004 getestet und dabei eine Treffergenauigkeit von bis zu 84,2 Prozent erreicht.

5 Nutzerbefragungen

Die Nutzerbefragungen im Bereich der digitalen Musik konzentrieren sich derzeit stark auf das Verhalten der Nutzer bei der Suche nach Musikinformationen (MIR), wie zum Beispiel die Befragung von S. Taheri-Panah und A. MacFarlane [10]. Nur wenige Studien beschäftigen sich mit Musikbibliotheken im Allgemeinen, bzw. der Organisation dieser im Speziellen. Die im Vorangehenden bereits angesprochene Untersuchung von realen CD-Sammlungen, durchgeführt von Sally Jo Cunningham, Matt Jones und Steve Jones, hat die Autoren zu einigen, teilweise bereits erwähnten Konzepten für Bibliotheksprogramme geführt. Eine weitere, nennenswerte Erkenntnis dieser Studie ist, dass das Albumcover einer CD von den Musiksammlern nicht nur als Zusatzinformation angesehen

wird, sondern ihnen auch beim Durchsuchen ihrer Sammlungen von Hilfe ist. Dies scheint nicht nur für die Sammler selber, sondern auch für Außenstehende von Nutzen zu sein, um sich einen Überblick über die Inhalte einer Sammlung zu machen. Die Integration von Cover-Art in iTunes hat also durchaus seine Berechtigung.

5.1 Welche Zusatzinformationen sind erwünscht?

Einige der Ergebnisse von MIR-Studien lassen sich auch auf die Konzeption von Bibliotheksprogrammen übertragen. So zeigt etwa die Studie "Music Information Needs, Uses, and Seeking Behaviours" von Jin Ha Lee und J. Stephen Downie [11], wie in Abbildung 3 dargestellt, nach welchen Musikinformatio- nien die Studienteilnehmer primär suchen. Während einige der genannten Informa- tionen, wie "Price of Item" oder "Sample tracks for listening" für ein Biblio- theksprogramm irrelevant sind, kann man doch anhand der Werte von "Lyrics" und "Artist information" durchaus ablesen, welche Metadaten für die Integra- tion in ein Sammlungsprogramm lohnenswert sind. Diese Werte werden auch in der "Digital Music Interaction Concepts: A User Study" von Fabio Vignoli [8] bestätigt, wobei dort das Albumcover als wesentlich wichtiger eingestuft wird.

Response	Positive	Negative	Don't know
	%	%	%
Music information			
Title of work(s)	90.1	7.4	2.5
Lyrics	81.0	15.4	3.6
Artist information	74.6	23.7	1.7
Sample tracks for listening	67.4	27.3	5.3
Track listing	60.7	33.8	5.5
Price of item	51.7	41.5	6.8
Information on genre	49.1	46.3	4.6
Review/Rating by others	47.1	47.3	5.6
Influences among artists	42.6	52.6	4.8
Background information (history, theory, etc.)	39.1	55.4	5.6
Information on different version(s) of work(s)	37.3	55.7	7.0
Artwork/Album cover	30.8	62.8	6.5
Links to related websites	29.7	62.2	8.0
Released date	21.5	71.2	7.3
Record label	15.0	77.9	7.0

Abbildung 3. Antworten zu "How likely are you to seek the following music information?".

Die Autoren unterscheiden dabei zwischen "Content" und "Context" Me- tadata. Unter Content ordnen sie "Musical" Metadata, z.B. Songtempo oder Songlänge, sowie "Bibliographic" Metadata, also traditionell beschreibende In- formation wie Titel und Interpret. Context besteht dagegen aus der "Relational" Metadata - Genre und Ähnlichkeitsindikationen, sowie "Associative" Metadata, also die Verwendung von Musik in anderen Arbeiten und Medien, wie etwa im

Fernsehen oder in Filmen. Sie bewerten dabei die "Relational Metadata" am höchsten, da in der Befragung nach dem Suchverhalten ihrer Studienteilnehmer das "Genre" (62,7%) sowie "Similar artist(s)" und "Similar music" (59,3% und 54,2%) häufig als Anhaltspunkte genannt werden. Auch die Notwendigkeit von "Associative Metadata" lässt sich aus den Antworten auf die Frage, was eine Musikinformationssuche ausgelöst hat, ablesen. Hier wurde "Radio show" (81,2%), "TV show, movier or animation" (80,8%) und "Advertisement or commercial" (64,2%) oft genannt [11].

5.2 Musikauswahl nach Stimmung

Fabio Vignoli lässt in seiner Studie [8] eine Gruppe von Befragten ein Bibliotheksprogramm namens "Moodlogic" testen und vergleicht deren Aussagen anschließend mit einer Kontrollgruppe. Moodlogic gibt dem Nutzer die Möglichkeit, Musik nach Stimmung, Tempo und Erscheinungsjahr auszuwählen und eine Playlist im Sinne von "Query-by-example" anhand eines Beispielliedes zu erstellen. Die Klassifikation der Stimmung wurde von der Community des Moodlogic-Programmes manuell erstellt. Die Ergebnisse der Befragung zeigen, anders als vielleicht zu vermuten ist, dass die Benutzer von Moodlogic die Musiksuche nach Stimmung, Tempo und Erscheinungsjahr allesamt negativer bewerten als die Kontrollgruppe. Allerdings wird die Möglichkeit der "Query-by-example"-Suche sowohl auf einzelne Lieder, wie auch auf Interpreten bezogen von beiden Gruppen sehr positiv bewertet. Dies, zusammen mit der Tatsache, dass in den Leitfadenterviews im Vorfeld dieser Studie die Attribute Stimmung, Situation und Aktivität häufig im Zusammenhang mit Musikselektion und Musiksuche genannt wurden, lässt zumindest vermuten, dass eine Möglichkeit der Musikauswahloption nach Stimmung gewünscht wird. Vielleicht lassen sich aber Stimmungen nicht immer so leicht charakterisieren und bedeuten auch nicht für jeden das gleiche, weswegen eine "Query-by-example"-Auswahl von Musik für eine bestimmte Situation/Stimmung befürwortet wird [8].

6 Visualisierung und Interfaces

Bibliotheksprogramme, die in der Visualisierung der Musiksammlungen auf textuellen Listen basieren, geben dem Nutzer unter Umständen nicht genug Übersicht über seine Sammlung, um diese effizient zu organisieren. Neuartige Visualisierungen können dabei helfen, dem Musiksammler einen besseren Gesamteindruck seiner Sammlung zu gewähren und die Erstellung und Organisation von Playlisten intuitiver zu gestalten. Genau diese beiden Ziele haben sich Marc Torrens, Patrick Hertzog und Josep-Lluís Arcos bei der im Nachfolgenden beschriebenen "Disc-Visualisierung" gesetzt [4].

6.1 Alternative Visualisierungen von Bibliotheken

Wie in Abbildung 4 zu sehen ist wird eine Disc in verschiedene Sektoren unterteilt, welche jeweils ein Genre der Bibliothek repräsentieren und dementsprechende Größenverhältnisse haben. Diese Sektoren sind wieder in gleicher

Weise in Subsektoren aufgeteilt, welche Interpreten in diesen Genres darstellen. Der Radius der Disc dient als Zeitskala, wobei Lieder mit einem früheren Entstehungsjahr in der Mitte und neuere Songs am Rand plaziert sind. Einzelne Tracks sind als Punkte auf der Disc dargestellt und zu Alben gruppiert, weswegen der Eindruck von Bögen entsteht. Soweit sortiert also die Visualisierung nach Genre, Interpret, Entstehungsjahr und Album. Der Nutzer darf nun eine fünfte quantitative Dimension bestimmen, z.B. Playcount (Dt. Abspielhäufigkeit), zuletzt abgespielt oder eine zuvor vergebene Wertung für MP3s. Diese Dimension wird farblich auf der Disc hervorgehoben (schwarz-hellblau). In Abbildung 4 ist auch zu sehen, wie in dieser Darstellung Playlisten visualisiert werden. Playlisten, welche keiner Gruppierungslogik folgen, werden mit geometrischen Figuren gekennzeichnet, etwa große rote Kreuze für die Playlist "25 last played". Die anderen Playlisten sind in der Visualisierung Rot hinterlegt, da ihre Gruppierungslogik dies zulässt - "60s & 70s music" ist als Kreis zu sehen.

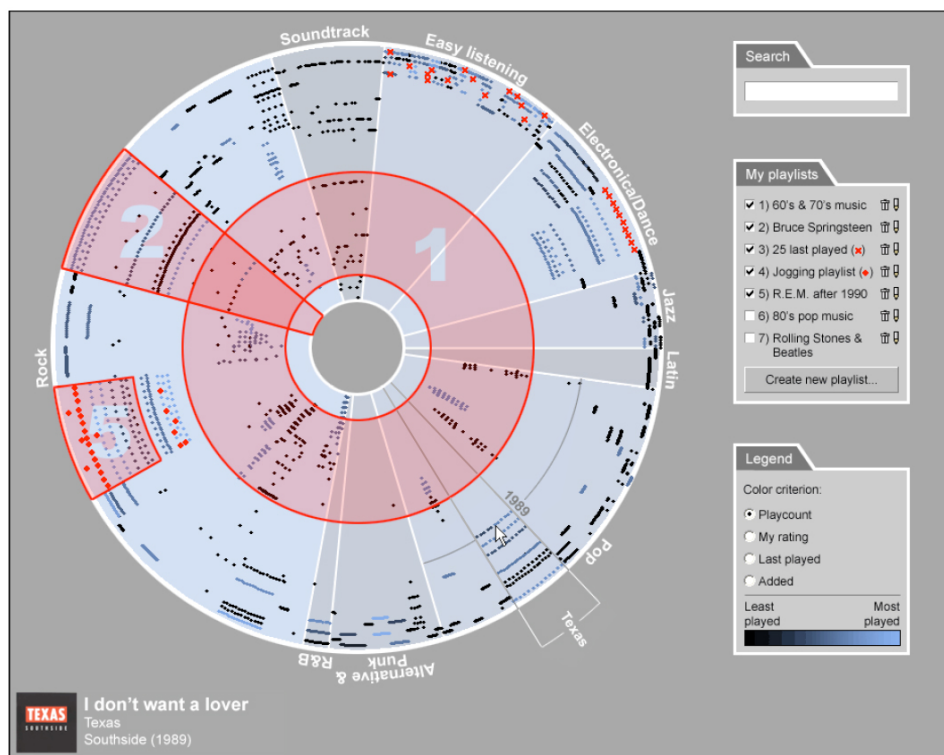


Abbildung 4. Disc-Visualisierung mit Playlists und Smart-Playlists.

Die Metadaten eines Tracks in der Bibliothek werden in der rechten unteren Ecke angezeigt, wenn man sich mit der Maus über diesem Track befindet. Zeitgleich ist der Interpret dieser Sektion hervorgehoben und das Entstehungsjahr des Tracks als Ring angezeigt. Der Nutzer kann in die Visualisierung hineinzoomen, wobei eine neue Disc entsteht, welche dann nur noch ein Genre enthält. Sektoren repräsentieren in diesem Fall nur noch Interpreten. In der letzte Zoomstufe entsteht eine Disc ohne Sektoren, welche dann sämtliche Lieder eines

Interpreten nach Alben geordnet enthält. Darüber hinaus bietet das Programm graphische Hilfen zum Bearbeiten von Playlists an. Zum Beispiel können zwei Playlists vereinigt oder ihre Schnittmenge verwendet werden, wobei die resultierende Playliste sofort graphisch beurteilbar ist.

6.2 Interfaces für Portable Devices

Durch die Fortschritte im Bereich der Unterhaltungselektronik ist es möglich geworden, mehr und mehr Musik auf kleine portable Geräte zu speichern. Die neuesten Geräte haben immense Speicherkapazitäten und sind dabei kleiner als ein altmodischer Walkman. Um das Durchsuchen der Musiksammlungen auf diesen Geräten zu erleichtern, haben die meisten Hersteller eine Patentlösung, wie zum Beispiel das "Touch-Wheel" von iPod oder den "Super-Scroll" von Philips entwickelt, mit denen man die Listen-basierten Darstellungen auf den Geräten schnell durchblättern kann. Ein Überblick über die Sammlung wird dadurch aber leider nicht gegeben, weswegen es für den Nutzer schwer ist, ein Musikstück zu finden, dessen Titel er vergessen hat. Rob van Gulik, Fabio Vignoli und Huub van de Wetering stellen mit ihrer "Artist-Map" [9] ein Interface vor, das es dem Nutzer ermöglichen soll, eben diesen Überblick auf einem Display mit nur 10 cm Diagonale zu erhalten.

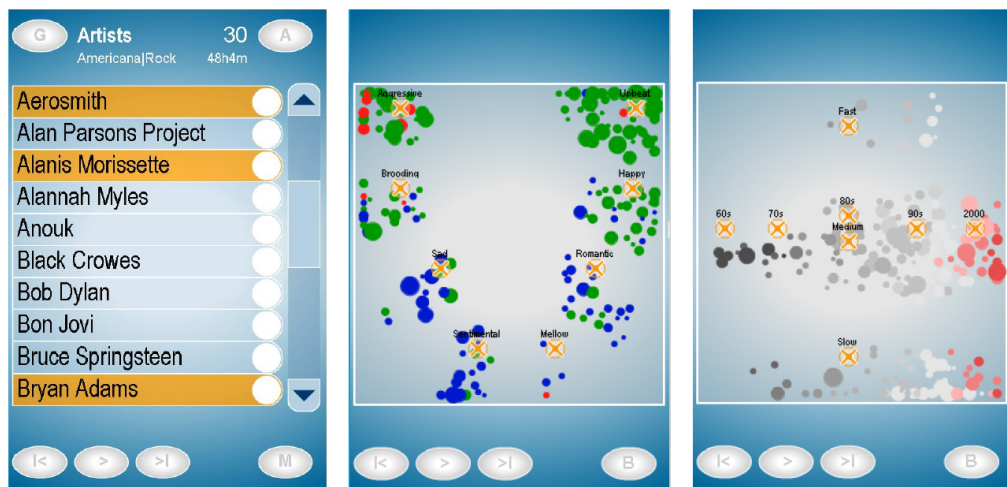


Abbildung 5. Die "Artist-Map"-Anwendung mit ihren unterschiedlichen Visualisierungen.

Um an die benötigte Metainformation für ihre Artist-Map zu gelangen verwenden die Autoren alle im Vorangehenden beschriebenen Wege. Über Online-Services gelangen sie an Interpreten- und Albumnamen, sowie Songtitel und Entstehungsjahr, über Soundanalyse versuchen sie Tempo und Genre herauszufiltern und über die Community-Klassifizierung von "Moodlogic" versuchen sie an Stimmungs-Einstufungen für Lieder zu gelangen. Das Bibliotheksprogramm selber hat dann neben einer traditionellen, Listen-basierten Darstellung der Musiksammlung auch eine Visualisierung über die "Artist-Map".

Auf ihr werden die Interpreten der Sammlung als Punkte auf einer Landkarte dargestellt. Desweiteren befinden sich auf der Landkarte Magneten, die bestimmten Attributen wie Stimmung, Tempo oder Entstehungsjahr entsprechen. Die Interpreten sind dann abhängig von ihrer Zugehörigkeit zu bestimmten Magneten auf der Karte plziert und zur besseren Übersicht auch farblich gekennzeichnet. Abbildung 5 zeigt sowohl die Listen-basierte Darstellung, als auch eine eindimensionale, nach Stimmung sortierte und eine zweidimensionale, nach Tempo und Entstehungsjahr sortierte Artist-Map. Anzahl, Attribut und Plazierung der Magneten kann der Nutzer frei wählen. Hat er in einen Bereich der Karte hineingezoomt und sich für einen oder mehrere Interpreten entschieden, so ändert sich das Interface wieder auf die Listen-basierte Darstellung, in der dann einzelne Musiktitel zum Abspielen ausgewählt werden können. Der Umstand, dass nicht die einzelnen Musikstücke, sondern die in der Anzahl deutlich geringeren Interpreten auf der Karte dargestellt werden, trägt dabei der Tatsache Rechnung, dass das Display eines Portable-Devices doch erheblich kleiner ist als das eines Monitors.

7 Fazit

Um einer immer größer werdenden Zahl von Sammlern digitaler Musik die Organisation ihrer immer größer werdenden Sammlungen zu erleichtern, ist eine effiziente und umfangreiche Extrahierung und Verifikation von Metadaten unabdingbar. Online-Dienste wie Free-DB und MusicBrainz leisten hierbei hervorragende Arbeit und verfügen heutzutage über sehr detaillierte Datenbanken. Da diese Datenbanken, um detailliert und umfassend zu bleiben, jedoch auf die Informationsgabe der Nutzer angewiesen sind, bleibt zu hoffen, dass nicht zu viele von ihnen den Weg der Kommerzialisierung beschreiten und kostenpflichtig werden, wie zum Beispiel CDDB. Das Genre stellt eine der wichtigsten Metainformationen für das Suchen in Musiksammlungen dar, da es dem Nutzer intuitiv aus Musikläden etc. vertraut ist und ihm die Informationslast abnimmt. Er kann dadurch zu Musik gelangen, die seinem derzeitigen Hörwunsch entspricht, ohne einen spezifischen Titel, Interpret oder Albumnamen zu wissen. Die Methoden der automatischen Genre-Extraktion via Sound-Analyse sind heute schon nahezu perfekt. Mit einer Treffergenauigkeit von rund 80% ist zum Beispiel der Rhythm-Pattern-Algorithmus von Rauber et al. vermutlich den meisten laienhaften Musikhörern weit überlegen [1].

Die mittlerweile recht große Palette an weiterführender Metainformation, welche dem Musiksammler heute zur Verfügung steht, ist erfreulich und hilft dem Bibliotheksbesitzer die Interaktion mit seiner Sammlung zu vertiefen. Es steht zu vermuten, dass dieser Informationsstrom mit zunehmenden Speicherkapazitäten und größer werdendem Vernetzungsgrad anschwellen wird. Nicht zuletzt deswegen, weil im Konkurrenzkampf kostenpflichtiger Bibliotheksprogramme dieser Mehrwert den kommerziellen Ausschlag geben kann. Die Umsetzung einer Visualisierung, die von der traditionellen Textbasiertheit weggeht, in einem der markführenden Bibliotheksprogramme wie iTunes, steht leider noch aus. Da in diesem Bereich jedoch schon diverse Konzepte entwickelt wurden, welche dem Nutzer eine bessere Übersicht über seine Sammlung gewähren und ein intuiti-

veres Browsen und Erstellen von Playlisten ermöglichen, bleibt dies wohl auch nur eine Frage der Zeit. iTunes geht mit seiner Smart-Playlist einen wichtigen Schritt in Richtung Sensibilisierung des Bibliotheksprogrammes auf die Hörgewohnheiten des Nutzers. Um diese Entwicklung voran zu treiben, wären intensivere Nutzerbefragungen auf diesem Gebiet sinnvoll.

Bei der Implementation aller dieser Systeme sollte jedoch die Leichtigkeit der Benutzung immer vorrangig sein. Eine Software, die zu umständlich in der Nutzung ist, macht die Organisation einer Musiksammlung nicht einfacher, sondern schwerer. Ein Umstand, der Sally Jo Cunningham et al. zu dem Schluss bringen: "Learning to use, and using, the system should not interfere with the enjoyment of a music collection" [3].

Literatur

- [1] Mahia Parparita, Brian Kernigham: *Applications of Web Services to Digital Music Libraries*, May 2003
- [2] James Howison, Abby Goodrum: *Why can't I manage academic papers like MP3s? The evolution and intent of Metadata standards*, June 2004
- [3] Sally Jo Cunningham, Matt Jones, Steve Jones: *Organizing Digital Music for Use: An Examination of Personal Music Collections*, 2004
- [4] Marc Torrens, Patrick Hertzog, Josep-Lluís Arcos: *Visualizing and Exploring Personal Music Libraries*, 2004
- [5] Andreas Rauber, Markus Fürwirth: *Automatically Analyzing and Organizing Music Archives*, 2001
- [6] Andreas Rauber, Elias Pampalk, Dieter Merkl: *Content-based Music Indexing and Organization*, August 2002
- [7] Andreas Rauber, Thomas Lidy, Robert Neumayer: *Content-based Organization of Digital Audio Collections*, 2004
- [8] Fabio Vignoli: *Digital Music Interaction Concepts: A User Study*, 2004
- [9] Robert van Gulik, Fabio Vignoli, Huub van de Wetering: *Mapping Music in the Palm of your Hand, Explore and Discover your Collection*, 2004
- [10] S. Taheri-Panah, A MacFarlane: *Music Information Retrieval systems: why do individuals use them and what are their needs?*, 2004
- [11] Jin Ha Lee, J. Stephen Downie: *Survey of Music information Needs, Uses, and Seeking Behaviours: Preliminary Findings*, 2004
- [12] <http://www.id3.org/> : *ID3v2 - The audience is informed*, Updated 2005
- [13] <http://de.wikipedia.org/wiki/ID3-Tag> : *Wikipedia - ID3-Tag*, Updated 2005
- [14] <http://de.wikipedia.org/wiki/MP3> : *Wikipedia - MP3*, Updated 2005
- [15] Steffen Pauws: *CubyHum: A Fully Operational Query by Humming System*, 2002
- [16] Elias Pampalk: *Islands of Music, Analysis, Organization, and Visualization of Music Archives*, 2001
- [17] George Tzanetakis, Perry Cook: *Musical Genre Classification of Audio Signals*, July 2002
- [18] Mike Slocombe: *MP3 Player Sales Set To Nearly Quadruple By 2009*, 17 März 2005

- [19] Mike Slocombe: *CES 2006 Starts: MP3 Player Sales To Soar 200%*, 05 Jan 2006
- [20] Janko Röttgers: *iPod-Konkurrenten kämpfen um Platz zwei*, 26. Juli 2005
- [21] Tony Smith: *Apple iTunes sales tally passes 350 million*, 14. April 2005
- [22] osViews.com: *The History of Apple's iPod*, 03. Mai 2005