Music Information Retrieval: Recent Developments and Applications

Markus Schedl
Johannes Kepler University Linz, Austria
markus.schedl@jku.at

Emilia Gómez
Universitat Pompeu Fabra, Barcelona, Spain
emilia.gomez@upf.edu

Julián Urbano
Universitat Pompeu Fabra, Barcelona, Spain
julian.urbano@upf.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Information Retrieval

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

M. Schedl , E. Gómez and J. Urbano. *Music Information Retrieval: Recent Developments and Applications*. Foundations and Trends® in Information Retrieval, vol. 8, no. 2-3, pp. 127–261, 2014.

This Foundations and Trends® issue was typeset in LaTeX using a class file designed by Neal Parikh. Printed on acid-free paper.

ISBN: 978-1-60198-807-2

Foundations and Trends® in Information Retrieval
Volume 8, Issue 2-3, 2014

Editorial Scope

Topics

Foundations and Trends® in Information Retrieval publishes survey and tutorial articles in the following topics:

- Applications of IR
- Architectures for IR
- Collaborative filtering and recommender systems
- Cross-lingual and multilingual IR
- Distributed IR and federated search
- Evaluation issues and test collections for IR
- Formal models and language models for IR
- IR on mobile platforms
- Indexing and retrieval of structured documents
- Information categorization and clustering
- Information extraction
- Information filtering and routing
- Metasearch, rank aggregation, and data fusion
- Natural language processing for IR
- Performance issues for IR systems, including algorithms, data structures, optimization techniques, and scalability
- Question answering
- Summarization of single documents, multiple documents, and corpora
- Text mining
- Topic detection and tracking
- Usability, interactivity, and visualization issues in IR
- User modelling and user studies for IR
- Web search

Information for Librarians

Foundations and Trends® in Information Retrieval, 2014, Volume 8, 5 issues. ISSN paper version 1554-0669. ISSN online version 1554-0677. Also available as a combined paper and online subscription.

Foundations and Trends® in Information Retrieval
Vol. 8, No. 2-3 (2014) 127–261

DOI: 10.1561/1500000042

Music Information Retrieval: Recent Developments and Applications

Markus Schedl
Johannes Kepler University Linz, Austria
markus.schedl@jku.at

Emilia Gómez
Universitat Pompeu Fabra, Barcelona, Spain
emilia.gomez@upf.edu

Julián Urbano
Universitat Pompeu Fabra, Barcelona, Spain
julian.urbano@upf.edu

Contents

Abstract

We provide a survey of the field of Music Information Retrieval (MIR), in particular paying attention to latest developments, such as semantic auto-tagging and user-centric retrieval and recommendation approaches. We first elaborate on well-established and proven methods for feature extraction and music indexing, from both the audio signal and contextual data sources about music items, such as web pages or collaborative tags. These in turn enable a wide variety of music retrieval tasks, such as semantic music search or music identification ("query by example"). Subsequently, we review current work on user analysis and modeling in the context of music recommendation and retrieval, addressing the recent trend towards user-centric and adaptive approaches and systems. A discussion follows about the important aspect of how various MIR approaches to different problems are evaluated and compared. Eventually, a discussion about the major open challenges concludes the survey.

M. Schedl , E. Gómez and J. Urbano. *Music Information Retrieval: Recent Developments and Applications*. Foundations and Trends® in Information Retrieval, vol. 8, no. 2-3, pp. 127–261, 2014.

DOI: 10.1561/1500000042.

1

Introduction to Music Information Retrieval

1.1 Motivation

Music is a pervasive topic in our society as almost everyone enjoys listening to it and many also create. Broadly speaking, the research field of Music Information Retrieval (MIR) is foremost concerned with the *extraction and inference of meaningful features from music* (from the audio signal, symbolic representation or external sources such as web pages), *indexing of music* using these features, and the development of different *search and retrieval* schemes (for instance, content-based search, music recommendation systems, or user interfaces for browsing large music collections), as defined by Downie [52]. As a consequence, MIR aims at making the world's vast store of music available to individuals [52]. To this end, different representations of music-related subjects (e.g., songwriters, composers, performers, consumer) and items (music pieces, albums, video clips, etc.) are considered.

Given the relevance of music in our society, it comes as a surprise that the research field of MIR is a relatively young one, having its origin less than two decades ago. However, since then MIR has experienced a constant upward trend as a research field. Some of the most important reasons for its success are (i) the development of audio compression techniques in the late 1990s, (ii) increasing computing power of personal computers, which in turn enabled users and applications to extract music features in a reasonable time, (iii) the widespread availability of mobile music players, and more recently (iv) the emergence of music streaming services such as *Spotify*[1], *Grooveshark*[2], *Rdio*[3] or *Deezer*[4], to name a few, which promise unlimited music consumption every time and everywhere.

1.2 History and evolution

Whereas early MIR research focused on working with symbolic representations of music pieces (i.e. a structured, digital representation of musical scores such as MIDI), increased computing power enabled the application of the full armory of signal processing techniques directly to the music audio signal during the early 2000s. It allowed the processing not only of music scores (mainly available for Western Classical music) but all kinds of recorded music, by deriving different music qualities (e.g. rhythm, timbre, melody or harmony) from the audio signal itself, which is still a frequently pursued endeavor in today's MIR research as stated by Casey et al. [28].

In addition, many important attributes of music (e.g. genre) are related not only to music content, but also to contextual/cultural aspects that can be modeled from user-generated information available for instance on the Internet. To this end, since the mid-2000s different data sources have been analyzed and exploited: web pages, microblogging messages from *Twitter*[5], images of album covers, collaboratively generated tags and data from games with a purpose.

Recently and in line with other related disciplines, MIR is seeing a shift — away from system-centric towards user-centric designs, both in models and evaluation procedures as mentioned by different authors such as Casey et al. [28] and Schedl et al. [241]. In the case of

[1] http://www.spotify.com
[2] http://grooveshark.com/
[3] http://www.rdio.com/
[4] http://www.deezer.com
[5] http://www.twitter.com

user-centric models, aspects such as serendipity (measuring how positively surprising a recommendation is), novelty, hotness, or location- and time-awareness have begun to be incorporated into models of users' individual music taste as well as into actual music retrieval and recommendation systems (for instance, in the work by Zhang et al. [307]).

As for evaluation, user-centric strategies aim at taking into account different factors in the perception of music qualities, in particular of music similarity. This is particularly important as the notions of music similarity and of music genre (the latter often being used as a proxy for the former) are ill-defined. In fact several authors such as Lippens et al. [157] or Seyerlehner [252] have shown that human agreement on which music pieces belong to a particular genre ranges only between 75% and 80%. Likewise, the agreement among humans on the similarity between two music pieces is also bounded at about 80% as stated in the literature [282, 230, 287, 112].

1.3 Music modalities and representations

Music is a highly multimodal human artifact. It can come as audio, symbolic representation (score), text (lyrics), image (photograph of a musician or album cover), gesture (performer) or even only a mental model of a particular tune. Usually, however, it is a mixture of these representations that form an individual's model of a music entity. In addition, as pointed out by Schedl et al. [230], human perception of music, and of music similarity in particular, is influenced by a wide variety of factors as diverse as lyrics, beat, perception of the performer by the user's friends, or current mental state of the user. Computational MIR approaches typically use features and create models to describe music by one or more of the following categories of music perception: *music content*, *music context*, *user properties*, and *user context*, as shown in Figure 1.1 and specified below.

From a general point of view, *music content* refers to aspects that are encoded in the audio signal, while *music context* comprises factors that cannot be extracted directly from the audio but are nevertheless related to the music item, artist, or performer. To give some exam-

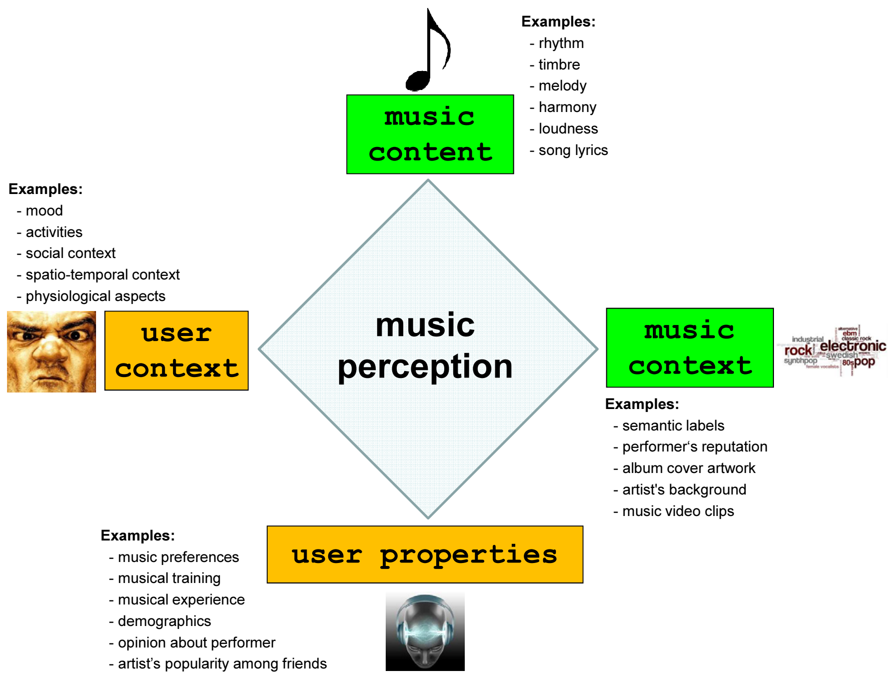

Figure 1.1: Categorization of perceptual music descriptors proposed in [230]

ples, rhythmic structure, melody, and timbre features belong to the former category, whereas information about an artist's cultural or political background, semantic labels, and album cover artwork belong to the latter. When focusing on the user, *user context* aspects represent dynamic and frequently changing factors, such as the user's current social context, activity, or emotion. In contrast, *user properties* refer to constant or only slowly changing characteristics of the user, such as her music taste or music education, but also the user's (or her friends') opinion towards a performer. The aspects belonging to user properties and user context can also be related to long-term and short-time interests or preferences. While user properties are tied to general, long-term goals, user context much stronger influences short-time listening needs.

Please note that there are interconnections between some features from different categories. For instance, aspects reflected in collaborative tags (e.g. musical genre) can be modeled by music content (e.g.

instrumentation) while some others (e.g. geographical location, influences) are linked to music context. Another example is semantic labels, which can be used to describe both the mood of a music piece and the emotion of a user as reviewed by Yang and Chen [305].

Ideally, music retrieval and recommendation approaches should incorporate aspects of several categories to overcome the "semantic gap", that is, the mismatch between machine-extractable music features and semantic descriptors that are meaningful to human music perception.

1.4 Applications

MIR as a research field is driven by a set of core applications that we present here from a user point of view.

1.4.1 Music retrieval

Music retrieval applications are intended to help users find music in large collections by a particular similarity criterion. Casey et al. [28] and Grosche et al. [89] propose a way to classify retrieval scenarios according to *specificity* (high specificity to identify a given audio signal and low to get statistically similar or categorically similar music pieces) and *granularity* or temporal scope (large granularity to retrieve complete music pieces and small granularity to locate specific time locations or fragments). Some of the most popular music retrieval tasks are summarized in the following, including pointers to respective scientific and industrial work.

Audio identification or *fingerprinting* is a retrieval scenario requiring high specificity and low granularity. The goal here is to retrieve or identify the same fragment of a given music recording with some robustness requirements (e.g. recording noise, coding). Well-known approaches such as the one proposed by Wang [297] have been integrated into commercially available systems, such as *Shazam*[6] (described in [297]), *Vericast*[7] or *Gracenote MusicID*[8]. Audio fingerprinting technolo-

[6] `http://www.shazam.com`
[7] `http://www.bmat.com/products/vericast/`
[8] `http://www.gracenote.com/music/recognition/`

gies are useful, for instance, to identify and distribute music royalties among music authors.

Audio alignment, matching or synchronization is a similar scenario of music retrieval where, in addition to identifying a given audio fragment, the aim is to locally link time positions from two music signals. Moreover, depending on the robustness of the audio features, one could also align different performances of the same piece. For instance, *MATCH* by Dixon and Widmer [48] and the system by Müller et al. [180] are able to align different versions of Classical music pieces by applying variants of the *Dynamic Time Warping* algorithm on sequences of features extracted from audio signals.

Cover song identification is a retrieval scenario that goes beyond the previous one (lower specificity level), as the goal here is to retrieve different versions of the same song, which may vary in many aspects such as instrumentation, key, harmony or structure. Systems for version identification, as reviewed by Serrà et al. [248], are mostly based on describing the melody or harmony of music signals and aligning these descriptors by local or global alignment methods. Web sites such as *The Covers Project*[9] are specialized in cover songs as a way to study musical influences and quotations.

In *Query by humming* and *query by tapping*, the goal is to retrieve music from a given melodic or rhythmic input (in audio or symbolic format) which is described in terms of features and is compared to the documents in a music collection. One of the first proposed systems is *MUSART* by Birmingham et al. [43]. Music collections for this task were traditionally built with music scores, user hummed or tapped queries –more recently with audio signals as in the system by Salamon et al. [218]. Commercial systems are also exploiting the idea of retrieving music by singing, humming or typing. One example is *SoundHound*[10], that matches users' hummed queries against a proprietary database of hummed songs.

The previously mentioned applications are based on the comparison of a target music signal against a database (also referred as *query by ex-*

[9] `http://www.coversproject.com/`
[10] `http://www.soundhound.com`

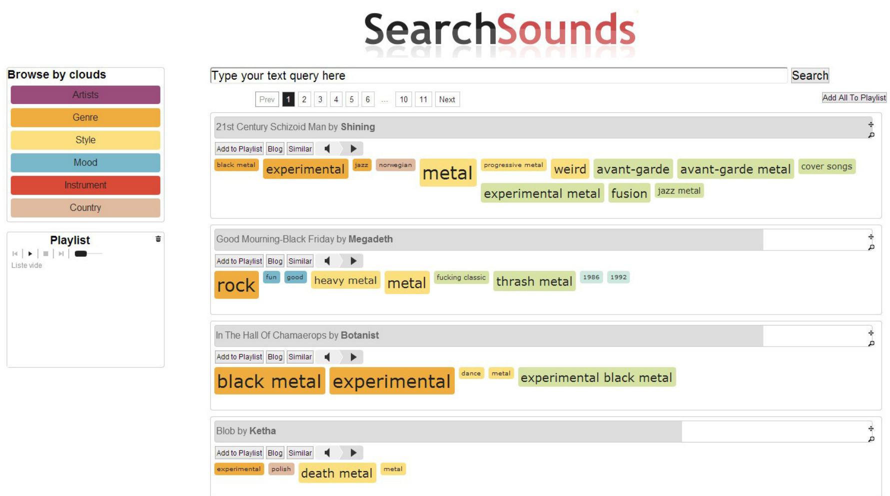

Figure 1.2: *SearchSounds* user interface for the query "metal".

ample), but users may want to find music fulfilling certain requirements (e.g. "give me songs with a tempo of 100 bpm or in C major") as stated by Isaacson [110]. In fact, humans mostly use *tags* or semantic descriptors (e.g. "happy" or "rock") to refer to music. *Semantic/tag-based or category-based retrieval* systems such as the ones proposed by Knees et al. [125] or Turnbull et al. [278] rely on methods for the estimation of semantic labels from music. This retrieval scenario is characterized by a low specificity and long-term granularity. An example of such semantic search engines is *SearchSounds* by Celma et al. [31, 266], which exploits user-generated content from music blogs to find music via arbitrary text queries such as "funky guitar riffs", expanding results with audio-based features. A screenshot of the user interface for the sample query "metal" can be seen in Figure 1.2. Another example is *Gedoodle* by Knees et al. [125], which is based on audio features and corresponding similarities enriched with editorial metadata (artist, album, and track names from ID3 tags) to gather related web pages. Both complementary pieces of information are then fused to map semantic user queries to actual music pieces. Figure 1.3 shows the results for the query "traditional irish".

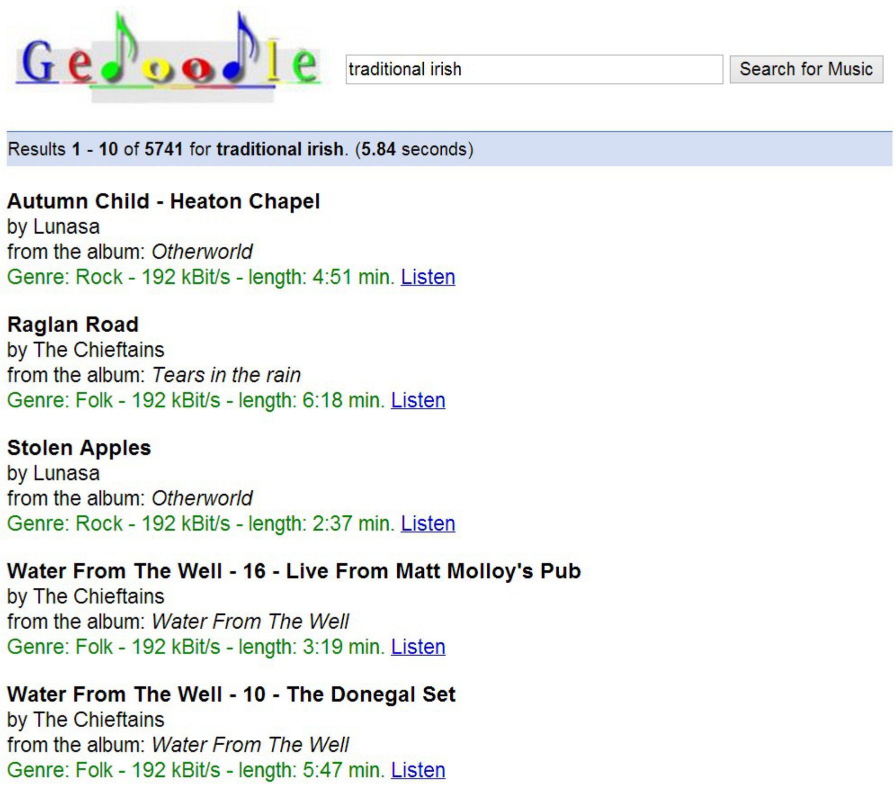

Figure 1.3: *Gedoodle* user interface for the query "traditional irish".

1.4.2 Music recommendation

Music recommendation systems typically propose a list of music pieces based on modeling the user's musical preferences. Ricci et al. [212] and Celma [30] state the main requirements of a recommender system in general and for music in particular: *accuracy* (recommendations should match one's musical preferences), *diversity* (as opposed to similarity, as users tend to be more satisfied with recommendations when they show a certain level of diversity), *transparency* (users trust systems when they understand why it recommends a music piece) and *serendipity* (a measure of "how surprising a recommendation is"). Well-known commercial systems are *Last.fm*[11], based on collaborative filtering, and *Pandora*[12], based on expert annotation of music pieces.

Recent methods proposed in the literature focus on user-aware, personalized, and multimodal recommendation. For example, Baltrunas et al. [7] propose their *InCarMusic* system for music recommendation in a car; Zhang et al. [307] present their *Auralist* music recommender with a special focus on serendipity; Schedl et al. [231, 238] investigate position- and location-aware music recommendation techniques based on microblogs; Forsblum et al. [70] propose a location-based recommender for serendipitous discovery of events at a music festival; Wang et al. [298] present a probabilistic model to integrate music content and user context features to satisfy user's short-term listening needs; Teng et al. [276] relate sensor features gathered from mobile devices with music listening events to improve mobile music recommendation.

1.4.3 Music playlist generation

Automatic music playlist generation, which is sometimes informally called "Automatic DJing", can be regarded as highly related to music recommendation. Its aim is to create an ordered list of results, such as music tracks or artists, to provide meaningful playlists enjoyable by the listener. This is also the main difference to general music recommendation, where the order in which the user listens to the recommended songs is assumed not to matter. Another difference between music recommendation and playlist generation is that the former typically aims at proposing new songs not known by the user, while the latter aims at reorganizing already known material.

A study conducted by Pohle et al. [206], in which humans evaluated the quality of automatically generated playlists, showed that similarity between consecutive tracks is an important requirement for a good playlist. Too much similarity between consecutive tracks, however, makes listeners feel bored by the playlist.

Schedl et al. [231] hence identify important requirements other than similarity: *familiarity/popularity* (all-time popularity of an artist or

[11] `http://www.lastfm.com`
[12] `http://www.pandora.com`

track), *hotness/trendiness* (amount of attention/buzz an artist currently receives), *recentness* (the amount of time passed since a track was released), and *novelty* (whether a track or artist is known by the user). These factors and some others contribute to a *serendipitous* listening experience, which means that the user is positively surprised because he encountered an unexpected, but interesting artist or song. More details as well as models for such serendipitous music retrieval systems can be found in [231] and in the work by Zhang et al. [307].

To give an example of an existing application that employs a content-based automatic playlist generation approach, Figure 1.4 depicts a screenshot of the *Intelligent iPod*[13] [246]. Audio features and corresponding similarities are directly extracted from the music collection residing on the mobile device. Based on these similarities, a playlist is created and visualized by means of a color stripe, where different colors correspond to different music styles, cf. (2) in Figure 1.4. The user can interact with the player with the scroll wheel to easily access the various music regions, cf. (4) in Figure 1.4.

Automatic playlist generation is also exploited in commercial products. To give an example, *YAMAHA BODiBEAT*[14] uses a set of body sensors to track one's workout and generate a playlist to match one's running pace.

1.4.4 Music browsing interfaces

Intelligent user interfaces that support the user in experiencing serendipitous listening encounters are becoming more and more important, in particular to deal with the abundance of music available to consumers today, for instance via music streaming services. These interfaces should hence support browsing through music collections in an intuitive way as well as retrieving specific items. In the following, we give a few examples of proposed interfaces of this kind.

The first one is the *nepTune*[15] interface proposed by Knees et al. [128], where music content features are extracted from a given mu-

[13] `http://www.cp.jku.at/projects/intelligent-ipod`
[14] `http://www.yamaha.com`
[15] `http://www.cp.jku.at/projects/neptune`

Figure 1.4: *Intelligent iPod* mobile browsing interface.

sic collection and then clustered. The resulting clusters are visualized by creating a virtual landscape of the music collection. The user can then navigate through this artificial landscape in a manner similar to a flight simulator game. Figure 1.5 shows screenshots of the *nepTune* interface. In both versions, the visualization is based on the metaphor of "Islands of Music" [193], according to which densely populated clusters of songs are visualized as mountains, whereas sparsely populated regions are visualized as beaches and oceans.

A similar three-dimensional browsing interface for music collections is presented by Lübbers and Jarke [161]. Unlike *nepTune*, which employs the "Islands of Music" metaphor, their system uses an inverse height map, by means of which clusters of music items are visualized as valleys separated by mountains corresponding to sparse regions. In addition, Lübbers and Jarke's interface supports user adaptation by providing means of deforming the landscape.

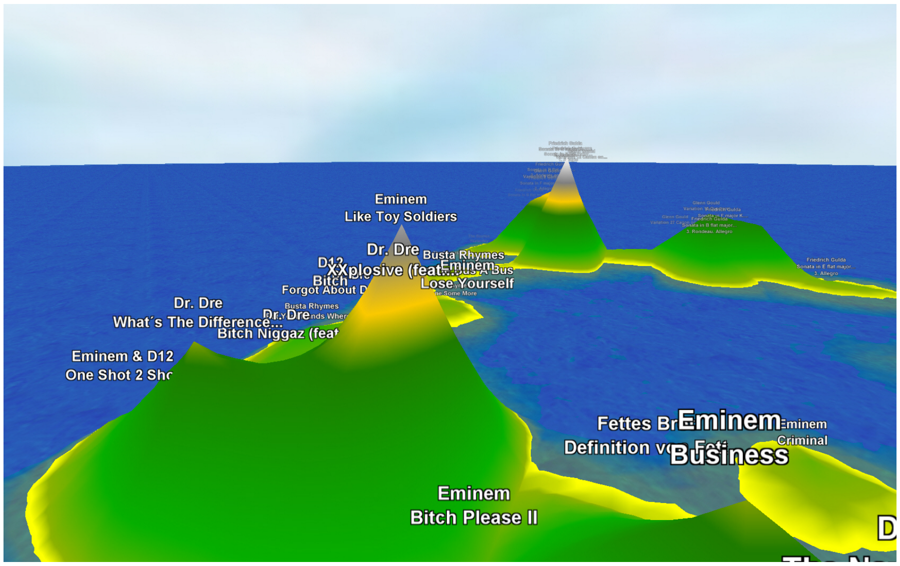

Figure 1.5: *nepTune* music browsing interface.

Musicream[16] by Goto and Goto [80] is another example of a user interface that fosters unexpected, serendipitous encounters with music, this time with the metaphor of a water tap. Figure 1.6 depicts a screenshot of the application. The interface includes a set of colored taps (in the top right of the figure), each corresponding to a different style of music. When the user decides to open the virtual handle, the respective tap creates a flow of songs. The user can then grab and play songs, or stick them together to create playlists (depicted on the left side of the figure). When creating playlists in this way, similar songs can be easily connected, whereas repellent forces are present between dissimilar songs, making it much harder to connect them.

Songrium[17] is a collection of web applications designed to enrich the music listening experience. It has been developed and is maintained by the National Institute of Advanced Industrial Science and Technology (AIST) in Japan. As illustrated by Hamasaki and Goto [90], *Songrium* offers various ways to browse music, for instance, via vi-

[16] `http://staff.aist.go.jp/m.goto/Musicream`
[17] `http://songrium.jp`

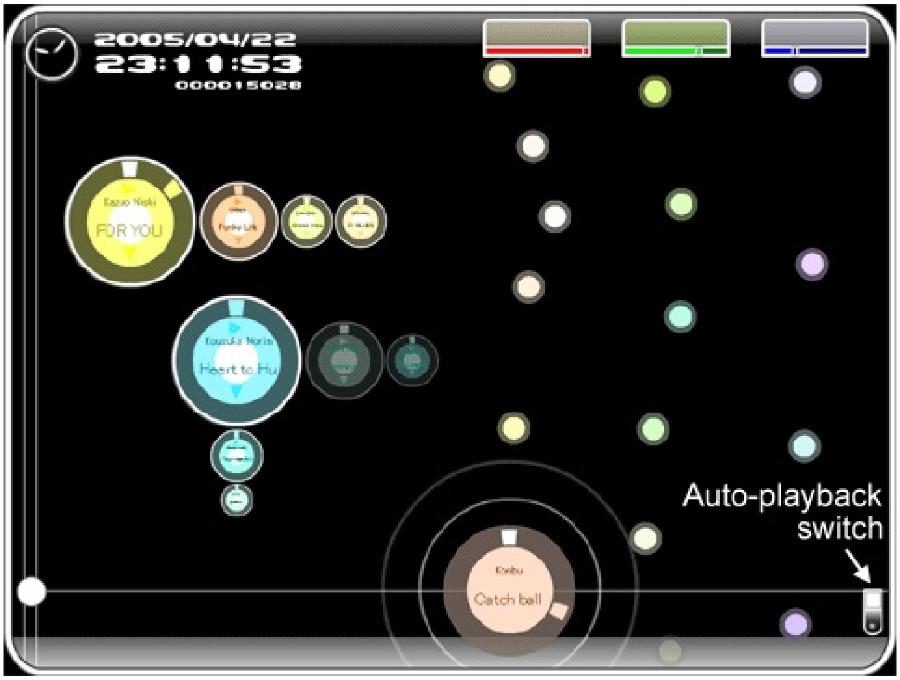

Figure 1.6: *Musicream* music browsing interface.

sualizing songs in a graph using audio-based similarity for placement ("Music Star Map"), via visualizing a song and its derivative works in a solar system-like structure ("Planet View"), or via exploring music by following directed edges between songs, which can be annotated by users ("Arrow View").

1.4.5 Beyond retrieval

MIR techniques are also exploited in other contexts, beyond the standard retrieval scenarios. One example is the computational music theory field, for which music content description techniques offer the possibility to perform comparative studies using large datasets and to formalize expert knowledge. In addition, music creation applications benefit from music retrieval techniques, for instance via "audio mosaicing", where a target music track is analyzed, its audio descriptors extracted for small fragments, and these fragments substituted with similar but novel fragments from a large music dataset. These applications are further reviewed in a recent "Roadmap for Music Information ReSearch" build by a community of researchers in the context of the MIReS project[18] [250].

1.5 Research topics and tasks

We have seen that research on MIR comprises a rich and diverse set of areas whose scope goes well beyond mere retrieval of documents, as pointed out by several authors such as Downie et al. [55, 20], Lee et al. [147, 148] and Bainbridge et al. [6]. MIR researchers have then been focusing on a set of concrete research tasks, which are the basis for final applications. Although most of the tasks will be reviewed within this manuscript, we already provide at this point an overview of some of the most important ones (including references) in Table 1.1.

A first group of topics are related to the extraction of meaningful features from music content and context. These features are then used to compute similarity between two musical pieces or to classify music pieces according to different criteria (e.g. mood, instrument, or genre). Features, similarity algorithms and classification methods are then tailored to different applications as described below.

1.6 Scope and related surveys

The field of MIR has undergone considerable changes during recent years. Dating back to 2006, Orio [186] presented one of the earliest survey articles on MIR, targeted at a general Information Retrieval audience who is already familiar with textual information. Orio does a great job in introducing music terminology and categories of music features that are important for retrieval. He further identifies different users of an MIR system and discusses their individual needs and requirements towards such systems. The challenges of extracting timbre, rhythm, and melody from audio and MIDI representations of music are discussed. To showcase a music search scenario, Orio discusses different

[18] `http://mires.eecs.qmul.ac.uk/`

ways of music retrieval via melody. He further addresses the topics of automatic playlist generation, of visualizing and browsing music collections, and of audio-based classification. Eventually, Orio concludes by reporting on early benchmarking activities to evaluate MIR tasks.
Although Orio's work gives a thorough introduction to MIR, many new research directions have emerged within the field since then. For instance, research on web-, social media-, and tag-based MIR could not be included in his survey. Also benchmarking activities in MIR were still in their fledgling stages at that time. Besides contextual MIR and evaluation, considerable progress has been made in the tasks listed in Table 1.1. Some of them even emerged only after the publication of [186]; for instance, auto-tagging or context-aware music retrieval.

Other related surveys include [28], where Casey et al. give an overview of the field of MIR from a signal processing perspective. They hence strongly focus on audio analysis and music content-based similarity and retrieval. In a more recent book chapter [227], Schedl gives an overview of music information extraction from the Web, covering the automatic extraction of song lyrics, members and instrumentation of bands, country of origin, and images of album cover artwork. In addition, different contextual approaches to estimate similarity between artists and between songs are reviewed. Knees and Schedl [127], give a survey of music similarity and recommendation methods that exploit contextual data sources. Celma's book [30] comprehensively addressed the problem of music recommendation from different perspectives, paying particular attention to the often neglected "long tail" of little-known music and how it can be made available to the interested music aficionado.

In contrast to these reviews, in this survey we (i) also discuss the very current topics of user-centric and contextual MIR, (ii) set the discussed techniques in a greater context, (iii) show applications and combinations of techniques, not only addressing single aspects of MIR such as music similarity, and (iv) take into account more recent work.

Given the focus of the survey at hand on recent developments in MIR, we decided to omit most work on symbolic (MIDI) music representations. Such work is already covered in detail in Orio's article

[186]. Furthermore, such work has been seeing a decreasing number of publications during the past few years. Another limitation of the scope is the focus on Western music, which is due to the fact that MIR research on music of other cultural areas is very sparse, as evidenced by Serra [249].

As MIR is a highly multidisciplinary research field, the annual "International Society for Music Information Retrieval" conference[19] (ISMIR) brings together researchers of fields as diverse as Electrical Engineering, Library Science, Psychology, Computer Science, Sociology, Mathematics, Music Theory, and Law. The series of ISMIR conferences are a good starting point to dig deeper into the topics covered in this survey. To explore particular topics or papers presented at ISMIR, the reader can use the *ISMIR Cloud Browser*[20] [88].

1.7 Organization of this survey

This survey is organized as follows. In Section 2 we give an overview of music content-based approaches to infer music descriptors. We discuss different categories of feature extractors (from low-level to semantically meaningful, high-level) and show how they can be used to infer music similarity and to classify music. In Section 3 we first discuss data sources belonging to the music context, such as web pages, microblogs, or music playlists. We then cover the tasks of extracting information about music entities from web sources and of music similarity computation for retrieval from contextual sources. Section 4 covers a very current topic in MIR research, i.e. the role of the user, which has been neglected for a long time in the community. We review ideas on how to model the user, highlight the crucial role the user has when elaborating MIR systems, and point to some of the few works that take the user context and the user properties into account. In Section 5 we give a comprehensive overview on evaluation initiatives in MIR and discuss their challenges. Section 6 summarizes this survey and highlights some of the grand challenges MIR is facing.

[19] `http://www.ismir.net`
[20] `http://dc.ofai.at/browser/all`

Table 1.1: Typical MIR subfields and tasks.

Task	**References**
FEATURE EXTRACTION	
Timbre description	Peeters et al. [200], Herrera et al. [99]
Music transcription and melody extraction	Klapuri & Davy [122], Salamon & Gómez[215], Hewlett & Selfridge-Field [103]
Onset detection, beat tracking, and tempo estimation	Bello et al. [10], Gouyon [83], McKinney & Breebaart [171]
Tonality estimation: chroma, chord, and key	Wakefield [296], Chew [34], Gómez [73], Papadopoulos & Peeters [197], Oudre et al. [188], Temperley [274]
Structural analysis, segmentation and summarization	Cooper & Foote [37], Peeters et al. [202], Chai [32]
SIMILARITY	
Similarity measurement	Bogdanov et al. [18], Slaney et al. [28], Schedl et al. [236, 228]
Cover song identification	Serra et al. [248], Bertin-Mahieux & Ellis [14]
Query by humming	Kosugi et al. [132], Salamon et al. [218], Dannenberg et al. [43]
CLASSIFICATION	
Emotion and mood recognition	Yang & Chen [304, 305], Laurier et al. [139]
Genre classification	Tzanetakis & Cook [281], Knees et al. [124]
Instrument classification	Herrera et al. [102]
Composer, artist and singer identification	Kim et al. [118]
Auto-tagging	Sordo [264], Coviello et al. [39], Miotto & Orio [173]
APPLICATIONS	
Audio fingerprinting	Wang [297], Cano et al. [24]
Content-based querying and retrieval	Slaney et al. [28]
Music recommendation	Celma [30], Zhang et al. [307], Kaminskas et al. [114]
Playlist generation	Pohle et al. [206], Reynolds et al. [211], Pampalk et al. [196], Aucouturier & Pachet [2]
Audio-to-score alignment and music synchronization	Dixon & Widmer [48], Müller et al. [180], Niedermayer [181]
Song/artist popularity estimation	Schedl et al. [237], Pachet & Roy [190] Koenigstein & Shavitt [130]
Music visualization	Müller & Jiang [179], Mardirossian & Chew [166], Cooper et al. [38], Foote [68], Gómez & Bonada [75]
Browsing user interfaces	Stober & Nürnberger [270], Leitich et al. [150], Lamere et al. [136], Pampalk & Goto [195]
Interfaces for music interaction	Steward & Sandler [268]
Personalized, context-aware and adaptive systems	Schedl & Schnitzer [238], Stober [269], Kaminskas et al. [114], Baltrunas et al. [7]

References

[1] Timothy G Armstrong, Alistair Moffat, William Webber, and Justin Zobel. Evaluatir: An online tool for evaluating and comparing ir systems. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 833–833, Boston, Massachusetts, USA, July 2009.

[2] Jean-Julien Aucouturier and François Pachet. Scaling Up Music Playlist Generation. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME 2002)*, pages 105–108, Lausanne, Switzerland, August 2002.

[3] Jean-Julien Aucouturier and François Pachet. Improving Timbre Similarity: How High is the Sky? *Journal of Negative Results in Speech and Audio Sciences*, 1(1), 2004.

[4] Claudio Baccigalupo, Enric Plaza, and Justin Donaldson. Uncovering Affinity of Artists to Multiple Genres from Social Behaviour Data. In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR'08)*, Philadelphia, PA, USA, September 14–18 2008.

[5] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval – the concepts and technology behind search.* Addison-Wesley, Pearson, Harlow, England, 2nd edition, 2011.

[6] David Bainbridge, Sally Jo Cunningham, and J Stephen Downie. How people describe their music information needs: A grounded theory analysis of music queries. In *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2003)*, pages 221–222, Baltimore, Maryland, USA, October 26–30 2003.

[7] Linas Baltrunas, Marius Kaminskas, Bernd Ludwig, Omar Moling, Francesco Ricci, Karl-Heinz Lüke, and Roland Schwaiger. InCarMusic: Context-Aware Music Recommendations in a Car. In *International Conference on Electronic Commerce and Web Technologies (EC-Web)*, Toulouse, France, Aug–Sep 2011.

[8] Eric Battenberg. Well-defined tasks and good datasets for mir: Ismir 2013 late-break. In *Proceedings of the International Society for Music Information Retrieval conference*, 2013.

[9] Stephan Baumann and Oliver Hummel. Using Cultural Metadata for Artist Recommendation. In *Proceedings of the 3rd International Conference on Web Delivering of Music (WEDELMUSIC 2003)*, Leeds, UK, September 15–17 2003.

[10] J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and Mark B. Sandler. A tutorial on onset detection in music signals. *Speech and Audio Processing, IEEE Transactions on*, 13(5):1035–1047, 2005.

[11] James Bennett and Stan Lanning. The netflix prize. In *Proceedings of KDD cup and workshop*, page 35, San Jose, California, USA, August 2007.

[12] Adam Berenzweig, Daniel P.W. Ellis, and Steve Lawrence. Anchor Space for Classification and Similarity Measurement of Music. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME 2003)*, Baltimore, Maryland, USA, July 2003. IEEE.

[13] Adam Berenzweig, Beth Logan, Daniel P.W. Ellis, and Brian Whitman. A Large-Scale Evaluation of Acoustic and Subjective Music Similarity Measures. In *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2003)*, Baltimore, MD, USA, October 26–30 2003.

[14] Thierry Bertin-Mahieux and Daniel P.W. Ellis. Large-Scale Cover Song Recognition Using the 2D Fourier Transform Magnitude. In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, Porto, Portugal, October 8-12 2012.

[15] Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The million song dataset. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, 2011.

[16] Jacob T. Biehl, Piotr D. Adamczyk, and Brian P. Bailey. DJogger: A Mobile Dynamic Music Device. In *CHI 2006: Extended Abstracts on Human Factors in Computing Systems*, pages 556–561, Montréal, Québec, Canada, 2006.

[17] D. Bogdanov, M. Haro, Ferdinand Fuhrmann, Anna Xambó, Emilia Gómez, and P. Herrera. Semantic audio content-based music recommendation and visualization based on user preference examples. *Information Processing & Management*, 49(1):13 – 33, 2013.

[18] D. Bogdanov, J. Serrà, N. Wack, and P. Herrera. From low-level to high-level: Comparative study of music similarity measures. In *IEEE International Symposium on Multimedia. Workshop on Advances in Music Information Research (AdMIRe)*, 2009.

[19] Dmitry Bogdanov, Joan Serrà, Nicolas Wack, Perfecto Herrera, and Xavier Serra. Unifying Low-Level and High-Level Music Similarity Measures . *IEEE Transactions on Multimedia*, 13(4):687–701, August 2011.

[20] Alain Bonardi. Ir for contemporary music: What the musicologist needs. In *Proceedings of the International Symposium on Music Information Retrieval (ISMIR 2000)*, Plymouth, MA, USA, October 2000.

[21] Martin Braschler and Carol Peters. Cross-language evaluation forum: Objectives, results, achievements. *Information retrieval*, 7(1-2):7–31, 2004.

[22] Eric Brill. A Simple Rule-based Part of Speech Tagger. In *Proceedings of the 3rd Conference on Applied Natural Language Processing (ANLC 1992)*, pages 152–155, Trento, Italy, March–April 1992.

[23] John Ashley Burgoyne, Bas de Haas, and Johan Pauwels. On comparative statistics for labelling tasks: what can we learn from MIREX ACE 2013? In *International Society for Music Information Retrieval Conference*, Taipei, Taiwan, October 2014.

[24] P. Cano, E. Batlle, E. Gómez, L. Gomes, and M. Bonnet. *Audio Fingerprinting Concepts and Applications.*, pages 233–245. Springer-Verlag, 2005.

[25] P. Cano, M. Koppenberger, S. Ferradans, A. Martinez, F. Gouyon, V. Sandvold, V. Tarasov, and N. Wack. MTG-DB: A Repository for Music Audio Processing. In *Proceedings of the 4th International Conference on Web Delivering of Music*, 2004.

[26] Pedro Cano, Emilia Gómez, Fabien Gouyon, Perfecto Herrera, Markus Koppenberger, Beesuan Ong, Xavier Serra, Sebastian Streich, and Nicolas Wack. Ismir 2004 audio description contest, 2006.

[27] Pedro Cano and Markus Koppenberger. The Emergence of Complex Network Patterns in Music Artist Networks. In *Proceedings of the 5th International Symposium on Music Information Retrieval (ISMIR 2004)*, pages 466–469, Barcelona, Spain, October 10–14 2004.

[28] Michael A. Casey, Remco Veltkamp, Masataka Goto, Marc Leman, Christophe Rhodes, and Malcolm Slaney. Content-Based Music Information Retrieval: Current Directions and Future Challenges. *Proceedings of the IEEE*, 96:668–696, April 2008.

[29] Toni Cebrián, Marc Planagumà, Paulo Villegas, and Xavier Amatriain. Music Recommendations with Temporal Context Awareness. In *Proceedings of the 4th ACM Conference on Recommender Systems*, Barcelona, Spain, 2010.

[30] Òscar Celma. *Music Recommendation and Discovery – The Long Tail, Long Fail, and Long Play in the Digital Music Space*. Springer, Berlin, Heidelberg, Germany, 2010.

[31] Òscar Celma, Pedro Cano, and Perfecto Herrera. SearchSounds: An Audio Crawler Focused on Weblogs. In *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR 2006)*, Victoria, Canada, October 8–12 2006.

[32] W. Chai. Semantic segmentation and summarization of music. *IEEE Signal Processing Magazine*, 23(2), 2006.

[33] Wei Chai. *Automated analysis of musical structure*. Doctoral dissertation, MIT, August 2005.

[34] Elaine Chew. *Towards a mathematical model of tonality*. PhD thesis, Massachusetts Institute of Technology, 2000.

[35] Ching-Hua Chuan and Elaine Chew. Polyphonic audio key finding using the spiral array ceg algorithm. In *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, pages 21–24. IEEE, 2005.

[36] William W. Cohen and Wei Fan. Web-Collaborative Filtering: Recommending Music by Crawling The Web. *WWW9 / Computer Networks*, 33(1–6):685–698, 2000.

[37] Matthew L. Cooper and Jonathan Foote. Automatic music summarization via similarity analysis. In *ISMIR*, 2002.

[38] Matthew L. Cooper, Jonathan Foote, Elias Pampalk, and George Tzanetakis. Visualization in audio-based music information retrieval. *Computer Music Journal*, 30(2):42–62, 2006.

[39] Emanuele Coviello, Antoni B. Chan, and Gert Lanckriet. Time Series Models for Semantic Music Annotation. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(5):1343–1359, July 2011.

[40] Sally Jo Cunningham, David Bainbridge, and J. Stephen Downie. The impact of mirex on scholarly research (2005 - 2010). In *Proceedings of the International Society for Music Information Retrieval conference*, pages 259–264, 2012.

[41] Sally Jo Cunningham and Jin Ha Lee. Influences of ismir and mirex research on technology patents. In *Proceedings of the International Society for Music Information Retrieval conference*, 2013.

[42] Stuart Cunningham, Stephen Caulder, and Vic Grout. Saturday Night or Fever? Context-Aware Music Playlists. In *Proceedings of the 3rd International Audio Mostly Conference of Sound in Motion*, October 2008.

[43] Roger B Dannenberg, William P Birmingham, Bryan Pardo, Ning Hu, Colin Meek, and George Tzanetakis. A comparative evaluation of search techniques for query-by-humming using the musart testbed. *Journal of the American Society for Information Science and Technology*, 58(5):687–701, 2007.

[44] Alain De Cheveigne. Pitch perception models. In *Pitch*, pages 169–233. Springer, 2005.

[45] Alain De Cheveigné and Hideki Kawahara. Yin, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111:1917, 2002.

[46] Scott Deerwester, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41:391–407, 1990.

[47] Peter Desain and Luke Windsor. *Rhythm perception and production*. Swets & Zeitlinger Publishers, 2000.

[48] Simon Dixon and Gerhard Widmer. Match: A music alignment tool chest. In *International Conference on Music Information Retrieval*, London, UK, 2005.

[49] Sandor Dornbush, Jesse English, Tim Oates, Zary Segall, and Anupam Joshi. XPod: A Human Activity Aware Learning Mobile Music Player. In *Proceedings of the Workshop on Ambient Intelligence, 20th International Joint Conference on Artificial Intelligence (IJCAI-2007)*, 2007.

[50] J. Stephen Downie. Interim Report on Establishing MIR/MDL Evaluation Frameworks: Commentary on Consensus Building. In *ISMIR Panel on Music Information Retrieval Evaluation Frameworks*, pages 43–44, 2002.

[51] J. Stephen Downie. *The MIR/MDL Evaluation Project White Paper Collection*. 3rd edition, 2003.

[52] J. Stephen Downie. Music Information Retrieval. *Annual Review of Information Science and Technology*, 37:295–340, 2003.

[53] J. Stephen Downie. The Scientific Evaluation of Music Information Retrieval Systems: Foundations and Future. *Computer Music Journal*, 28:12–23, June 2004.

[54] J Stephen Downie, Donald Byrd, and Tim Crawford. Ten years of ismir: Reflections on challenges and opportunities. In *International Conference on Music Information Retrieval (ISMIR'09)*, pages 13–18, 2009.

[55] J. Stephen Downie, Andreas F. Ehmann, Mert Bay, and M. Cameron Jones. *The Music Information Retrieval Evaluation eXchange: Some Observations and Insights*, pages 93–115. Springer, 2010.

[56] J Stephen Downie, Joe Futrelle, and David Tcheng. The international music information retrieval systems evaluation laboratory: Governance, access and security. In *Proceedings of 5th International Conference on Music Information Retrieval (ISMIR 2004)*, pages 9–14, Barcelona, Spain, October 10–14 2004.

[57] J. Stephen Downie, Kris West, Andreas F. Ehmann, and Emmanuel Vincent. The 2005 music information retrieval evaluation exchange (mirex 2005): Preliminary overview. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, pages 320–323, London, UK, September 11–15 2005.

[58] Karin Dressler. Sinusoidal extraction using an efficient implementation of a multi-resolution fft. In *Proc. of 9th Int. Conf. on Digital Audio Effects (DAFx-06)*, pages 247–252, 2006.

[59] Karin Dressler. Multiple fundamental frequency extraction for mirex 2012. *Eighth Music Information Retrieval Evaluation eXchange (MIREX)*, 2012.

[60] Gideon Dror, Noam Koenigstein, Yehuda Koren, and Markus Weimer. The yahoo! music dataset and kdd-cup'11. In *Proceedings of KDD cup and workshop*, San Diego, California, USA, August 2011.

[61] Andreas F. Ehmann, J. Stephen Downie, and M. Cameron Jones. The music information retrieval evaluation exchange ?o-it-yourself?web service. In *International Conference on Music Information Retrieval*, pages 323–324, Vienna, Austria, September 23-27 2007.

[62] Greg T. Elliott and Bill Tomlinson. Personalsoundtrack: Context-aware playlists that adapt to user pace. In *CHI 2006: Extended Abstracts on Human Factors in Computing Systems*, pages 736–741, Montr'eal, Qu'ebec, Canada, 2006.

[63] Daniel PW Ellis, Brian Whitman, Adam Berenzweig, and Steve Lawrence. The quest for ground truth in musical artist similarity. In *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR 2002)*, pages 170–177, Paris, France, October 13–17 2002.

[64] Daniel P.W. Ellis, Brian Whitman, Adam Berenzweig, and Steve Lawrence. The Quest For Ground Truth in Musical Artist Similarity. In *Proceedings of 3rd International Conference on Music Information Retrieval (ISMIR 2002)*, Paris, France, October 13–17 2002.

[65] A. Flexer. Statistical Evaluation of Music Information Retrieval Experiments. *Journal of New Music Research*, 35(2):113–120, June 2006.

[66] Arthur Flexer. On inter-rater agreement in audio music similarity. In *International Society for Music Information Retrieval Conference*, Taipei, Taiwan, October 2014.

[67] Arthur Flexer, Dominik Schnitzer, and Jan Schlüter. A MIREX meta-analysis of hubness in audio music similarity. In *ISMIR*, pages 175–180, 2012.

[68] J. Foote. Visualizing music and audio using self-similarity. In *ACM Multimedia*, pages 77–80, Orlando, USA, 1999.

[69] Jonathan Foote, Matthew L Cooper, and Unjung Nam. Audio retrieval by rhythmic similarity. In *ISMIR*, 2002.

[70] Andreas Forsblom, Petteri Nurmi, Pirkka Åman, and Lassi Liikkanen. Out of the bubble: Serendipitous even recommendations at an urban music festival. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces (IUI)*, pages 253–256, New York, NY, USA, 2012. ACM.

[71] Gijs Geleijnse, Markus Schedl, and Peter Knees. The Quest for Ground Truth in Musical Artist Tagging in the Social Web Era. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, Vienna, Austria, September 23–27 2007.

[72] Olivier Gillet and Gaël Richard. Enst-drums: an extensive audio-visual database for drum signals processing. In *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR 2006)*, pages 156–159, Victoria, Canada, October 8–12 2006.

[73] E. Gómez. Tonal description of polyphonic audio for music content processing. *INFORMS Journal on Computing, Special Cluster on Computation in Music*, 18, 2006.

[74] Emilia Gómez. *Tonal Description of Music Audio Signals*. PhD thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2006.

[75] Emilia Gómez and Jordi Bonada. Tonality visualization of polyphonic audio. In *Proceedings of International Computer Music Conference*. Citeseer, 2005.

[76] Emilia Gómez and Jordi Bonada. Towards computer-assisted flamenco transcription: An experimental comparison of automatic transcription algorithms as applied to a cappella singing. *Computer Music Journal*, 37(2):73–90, 2013.

[77] Emilia Gómez, Martín Haro, and Perfecto Herrera. Music and geography: Content description of musical audio from different parts of the world. In *ISMIR*, pages 753–758, 2009.

[78] Emilia Gómez, Anssi Klapuri, and Benoît Meudic. Melody description and extraction in the context of music content processing. *Journal of New Music Research*, 32(1):23–40, 2003.

[79] Masataka Goto. Grand Challenges in Music Information Research. In Meinard Müller, Masataka Goto, and Markus Schedl, editors, *Multimodal Music Processing*, volume 3 of *Dagstuhl Follow-Ups*, pages 217–225. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2012.

[80] Masataka Goto and Takayuki Goto. Musicream: New Music Playback Interface for Streaming, Sticking, Sorting, and Recalling Musical Pieces. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, London, UK, September 11–15 2005.

[81] Masataka Goto, Hiroki Hashiguchi, Takuichi Nishimura, and Ryuichi Oka. RWC music database: Popular, classical, and jazz music databases. In *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR 2002)*, pages 287–288, Paris, France, 2002.

[82] Masataka Goto, Hiroki Hashiguchi, Takuichi Nishimura, and Ryuichi Oka. RWC music database: Music genre database and musical instrument sound database. In *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2003)*, pages 229–230, Baltimore, Maryland, USA, October 26–30 2003.

[83] F. Gouyon. *Computational Rhythm Description*. VDM Verlag, 2008.

[84] Fabien Gouyon and Simon Dixon. A review of automatic rhythm description systems. *Computer music journal*, 29(1):34–54, 2005.

[85] Fabien Gouyon, Perfecto Herrera, Emilia Gomez, Pedro Cano, Jordi Bonada, Alex Loscos, and Xavier Amatriain; Xavier Serra. *Content Processing of Music Audio Signals*, chapter 3, pages 83–160. 978-3-8325-1600-0. 2008.

[86] Sten Govaerts and Erik Duval. A Web-based Approach to Determine the Origin of an Artist. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, Kobe, Japan, October 2009.

[87] Norbert Gövert and Gabriella Kazai. Overview of the initiative for the evaluation of xml retrieval (inex) 2002. In *Proceedings of the 1st Workshop of the INitiative for the Evaluation of XML Retrieval (INEX)*, pages 1–17, Dagstuhl, Germany, December 2002.

[88] Maarten Grachten, Markus Schedl, Tim Pohle, and Gerhard Widmer. The ISMIR Cloud: A Decade of ISMIR Conferences at Your Fingertips. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, Kobe, Japan, October 2009.

[89] Peter Grosche, Meinard Müller, and Joan Serrà. Audio Content-Based Music Retrieval. In Meinard Müller, Masataka Goto, and Markus Schedl, editors, *Multimodal Music Processing*, volume 3 of *Dagstuhl Follow-Ups*, pages 157–174. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2012.

[90] Masahiro Hamasaki and Masataka Goto. Songrium: A Music Browsing Assistance Service Based on Visualization of Massive Open Collaboration Within Music Content Creation Community. In *Proceedings of the 9th International Symposium on Open Collaboration*, WikiSym '13, pages 4:1–4:10, New York, NY, USA, 2013. ACM.

[91] Shuhei Hamawaki, Shintaro Funasawa, Jiro Katto, Hiromi Ishizaki, Keiichiro Hoashi, and Yasuhiro Takishima. Feature analysis and normalization approach for robust content-based music retrieval to encoded audio with different bit rates. In *Advances in Multimedia Modeling, International Multimedia Modeling Conference (MMM'08)*, page 298?309, 2008.

[92] Uri Hanani, Bracha Shapira, and Peretz Shoval. Information filtering: Overview of issues, research and systems. *User Modeling and User-Adapted Interaction*, 11(3):203–259, 2001.

[93] Donna K. Harman. Information retrieval evaluation. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 3(2):1–119, 2011.

[94] Christopher Harte. *Towards automatic extraction of harmony information from music signals*. PhD thesis, University of London, 2010.

[95] Christopher Harte, Mark Sandler, Samer Abdallah, and Emilia Gómez. Symbolic representation of musical chords: A proposed syntax for text annotations. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, pages 66–71, London, UK, September 11–15 2005.

[96] William M Hartmann. Pitch, periodicity, and auditory organization. *The Journal of the Acoustical Society of America*, 100:3491, 1996.

[97] David Hauger and Markus Schedl. Exploring Geospatial Music Listening Patterns in Microblog Data. In *Proceedings of the 10th International Workshop on Adaptive Multimedia Retrieval (AMR 2012)*, Copenhagen, Denmark, October 2012.

[98] David Hauger, Markus Schedl, Andrej Košir, and Marko Tkalčič. The Million Musical Tweets Dataset: What Can We Learn From Microblogs. In *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR 2013)*, Curitiba, Brazil, November 2013.

[99] P. Herrera, G. Peeters, and S. Dubnov. Automatic classification of musical instrument sounds. *Journal of New Music Research*, 32, 2003.

[100] P. Herrera, J. Serrà, C. Laurier, Enric Guaus, Emilia Gómez, and Xavier Serra. The discipline formerly known as mir. In *International Society for Music Information Retrieval (ISMIR) Conference, special session on The Future of MIR (fMIR)*, Kobe, Japan, 26/10/2009 2009.

[101] Perfecto Herrera, Zuriñe Resa, and Mohamed Sordo. Rocking around the clock eight days a week: an exploration of temporal patterns of music listening. In *1st Workshop On Music Recommendation And Discovery (WOMRAD), ACM RecSys, 2010, Barcelona, Spain*, 2010.

[102] Perfecto Herrera-Boyer, Anssi Klapuri, and Manuel Davy. Automatic classification of pitched musical instrument sounds. In Anssi Klapuri and Manuel Davy, editors, *Signal Processing Methods for Music Transcription*, pages 163–200. Springer US, 2006.

[103] Walter B Hewlett and Eleanor Selfridge-Field. *Melodic similarity: Concepts, procedures, and applications*, volume 11. The MIT Press, 1998.

[104] A. Holzapfel, M. E P Davies, J. R. Zapata, J.L. Oliveira, and F. Gouyon. On the automatic identification of difficult examples for beat tracking: Towards building new evaluation datasets. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 89–92, 2012.

[105] Andre Holzapfel, Matthew EP Davies, José R Zapata, João Lobato Oliveira, and Fabien Gouyon. Selective sampling for beat tracking evaluation. *Audio, Speech, and Language Processing, IEEE Transactions on*, 20(9):2539–2548, 2012.

[106] Helge Homburg, Ingo Mierswa, Bülent Möller, Katharina Morik, and Michael Wurst. A benchmark dataset for audio classification and clustering. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, pages 528–31, London, UK, September 11–15 2005.

[107] Xiao Hu, J. Stephen Downie, Kris West, and Andreas Ehmann. Mining Music Reviews: Promising Preliminary Results. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, London, UK, September 11–15 2005.

[108] Xiao Hu and Noriko Kando. User-centered Measures vs. System Effectiveness in Finding Similar Songs. In *Proc. ISMIR*, pages 331–336, Porto, Portugal, October 2012.

[109] Leonidas Ioannidis, Emilia Gómez, and Perfecto Herrera. Tonal-based retrieval of arabic and middle-east music by automatic makam description. In *Content-Based Multimedia Indexing (CBMI), 2011 9th International Workshop on*, pages 31–36. IEEE, 2011.

[110] E.J. Isaacson. Music IR for Music Theory. In *The MIR/MDL Evaluation Project White paper Collection*, pages 23–26, 2002.

[111] J. H. Jensen, M. G. Christensen, D. P. W. Ellis, and S. H. Jensen. Quantitative analysis of a common audio similarity measure. *IEEE Transactions on Audio, Speech, and Language Processing*, 17:693–703, 2009.

[112] M. Cameron Jones, J. Stephen Downie, and Andreas F. Ehmann. Human similarity judgments: Implications for the design of formal evaluations. In *International Conference on Music Information Retrieval*, pages 539–542, Vienna, Austria, September 23–27 2007.

[113] Marius Kaminskas and Francesco Ricci. Location-Adapted Music Recommendation Using Tags. In Joseph Konstan, Ricardo Conejo, José Marzo, and Nuria Oliver, editors, *User Modeling, Adaption and Personalization*, volume 6787 of *Lecture Notes in Computer Science*, pages 183–194. Springer Berlin / Heidelberg, 2011.

[114] Marius Kaminskas, Francesco Ricci, and Markus Schedl. Location-aware Music Recommendation Using Auto-Tagging and Hybrid Matching. In *Proceedings of the 7th ACM Conference on Recommender Systems (RecSys 2013)*, Hong Kong, China, October 2013.

[115] Noriko Kando, Kazuko Kuriyama, Toshihiko Nozue, Koji Eguchi, Hiroyuki Kato, and Souichiro Hidaka. Overview of ir tasks at the first ntcir workshop. In *Proceedings of the 1st NTCIR workshop on research in Japanese text retrieval and term recognition*, pages 11–44, Tokyo, Japan, September 1999.

[116] Joon Hee Kim, Brian Tomasik, and Douglas Turnbull. Using Artist Similarity to Propagate Semantic Information. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, Kobe, Japan, October 2009.

[117] Youngmoo E Kim, Erik Schmidt, and Lloyd Emelle. Moodswings: A collaborative game for music mood label collection. In *Proceedings of the 9th International Society for Music Information Retrieval Conference (ISMIR 2008)*, pages 231–236, Philadelphia, PA, USA, September 2008.

[118] Youngmoo E Kim and Brian Whitman. Singer identification in popular music recordings using voice coding features. In *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR)*, volume 13, page 17, 2002.

[119] Anssi Klapuri. Sound onset detection by applying psychoacoustic knowledge. In *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*, volume 6, pages 3089–3092. IEEE, 1999.

[120] Anssi Klapuri. Auditory model-based methods for multiple fundamental frequency estimation. In Klapuri and Davy [122], pages 229–265.

[121] Anssi Klapuri. Multiple fundamental frequency estimation by summing harmonic amplitudes. In *ISMIR*, pages 216–221, 2006.

[122] Anssi Klapuri and Manuel Davy, editors. *Signal Processing Methods for Music Transcription*. Springer, New York, 2006.

[123] Anssi P Klapuri. A perceptually motivated multiple-f0 estimation method. In *Applications of Signal Processing to Audio and Acoustics, 2005. IEEE Workshop on*, pages 291–294. IEEE, 2005.

[124] Peter Knees, Elias Pampalk, and Gerhard Widmer. Artist Classification with Web-based Data. In *Proceedings of the 5th International Symposium on Music Information Retrieval (ISMIR 2004)*, pages 517–524, Barcelona, Spain, October 10–14 2004.

[125] Peter Knees, Tim Pohle, Markus Schedl, and Gerhard Widmer. A Music Search Engine Built upon Audio-based and Web-based Similarity Measures. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2007)*, Amsterdam, the Netherlands, July 23–27 2007.

[126] Peter Knees and Markus Schedl. Towards Semantic Music Information Extraction from the Web Using Rule Patterns and Supervised Learning. In *Proceedings of the 2nd Workshop on Music Recommendation and Discovery (WOMRAD)*, Chicago, IL, USA, October 2011.

[127] Peter Knees and Markus Schedl. A survey of music similarity and recommendation from music context data. *Transactions on Multimedia Computing, Communications, and Applications*, 2013.

[128] Peter Knees, Markus Schedl, Tim Pohle, and Gerhard Widmer. An Innovative Three-Dimensional User Interface for Exploring Music Collections Enriched with Meta-Information from the Web. In *Proceedings of the 14th ACM International Conference on Multimedia (MM 2006)*, Santa Barbara, CA, USA, October 23–27 2006.

[129] Peter Knees and Gerhard Widmer. Searching for Music Using Natural Language Queries and Relevance Feedback. In *Proceedings of the 5th International Workshop on Adaptive Multimedia Retrieval (AMR'07)*, Paris, France, July 2007.

[130] Noam Koenigstein and Yuval Shavitt. Song Ranking Based on Piracy in Peer-to-Peer Networks. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, Kobe, Japan, October 2009.

[131] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix Factorization Techniques for Recommender Systems. *Computer*, 42:30–37, August 2009.

[132] Naoko Kosugi, Yuichi Nishihara, Tetsuo Sakata, Masashi Yamamuro, and Kazuhiko Kushima. A Practical Query-by-Humming System for a Large Music Database. In *Proceedings of the 8th ACM International Conference on Multimedia*, pages 333–342, Los Angeles, CA, USA, 2000.

[133] Andreas Krenmayer. Musikspezifische Informationsextraktion aus Webdokumenten. Master's thesis, Johannes Kepler University, Linz, Austria, August 2013.

[134] Carol L Krumhansl. *Cognitive foundations of musical pitch*, volume 17. Oxford University Press New York, 1990.

[135] Paul Lamere. Social Tagging and Music Information Retrieval. *Journal of New Music Research: Special Issue: From Genres to Tags – Music Information Retrieval in the Age of Social Tagging*, 37(2):101–114, 2008.

[136] Paul Lamere and Douglas Eck. Using 3D Visualizations to Explore and Discover Music. In *Proceedings of the 8th International Conference on Music Information Retrieval*, pages 173–174, Vienna, Austria, September 23-27 2007.

[137] M. Larson, M. Soleymani, P. Serdyukov, V. Murdock, and G.J.F. Jones, editors. *Working Notes Proceedings of the MediaEval 2010 Workshop*, 2010.

[138] Olivier Lartillot and Petri Toiviainen. A matlab toolbox for musical feature extraction from audio. In *International Conference on Digital Audio Effects*, pages 237–244, 2007.

[139] C. Laurier, O. Meyers, J. Serrà, M. Blech, P. Herrera, and X. Serra. Indexing music by mood: Design and integration of an automatic content-based annotator. *Multimedia Tools and Applications*, 48:161–184, 05/2010 2010. Springerlink link: http://www.springerlink.com/content/jj01750u20267426.

[140] E. Law, L. von Ahn, R. Dannenberg, and M. Crawford. Tagatune: A Game for Music and Sound Annotation. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, Vienna, Austria, September 2007.

[141] Edith Law and Luis von Ahn. Input-Agreement: A New Mechanism for Collecting Data Using Human Computation Games. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, CHI '09, pages 1197–1206, New York, NY, USA, 2009. ACM.

[142] Daniel D. Lee and H. Sebastian Seung. Learning the Parts of Objects by Non-negative Matrix Factorization. *Nature*, 401(6755):788–791, 1999.

[143] Jae Sik Lee and Jin Chun Lee. Context Awareness by Case-Based Reasoning in a Music Recommendation System. In Haruhisa Ichikawa, We-Duke Cho, Ichiro Satoh, and Hee Youn, editors, *Ubiquitous Computing Systems*, volume 4836 of *Lecture Notes in Computer Science*, pages 45–58. Springer Berlin / Heidelberg, 2007.

[144] J.H. Lee and N.M. Waterman. Understanding user requirements for music information services. In *International Society for Music Information Retrieval Conference*, pages 253–258, Porto, Portugal, October 2012.

[145] Jin Ha Lee. Crowdsourcing music similarity judgments using mechanical turk. In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, page 183?88, Utrecht, the Netherlands, August 2010.

[146] Jin Ha Lee and Sally Jo Cunningham. Toward an understanding of the history and impact of user studies in music information retrieval. *Journal of Intelligent Information Systems*, 41(3):499–521, 2013.

[147] Jin Ha Lee and J Stephen Downie. Survey of music information needs, uses, and seeking behaviours: Preliminary findings. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR 2004)*, pages 441–446, Barcelona, Spain, October 10–14 2004.

[148] Jin Ha Lee, J Stephen Downie, and Sally Jo Cunningham. Challenges in cross-cultural/multilingual music information seeking. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, pages 11–15, London, UK, September 11–15 2005.

[149] Jin Ha Lee and Xiao Hu. Generating ground truth for music mood classification using mechanical turk. In *Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries*, pages 129–138, 2012.

[150] Stefan Leitich and Martin Topf. Globe of music - music library visualization using geosom. In *Proceedings of the 8th International Conference on Music Information Retrieval*, pages 167–170, Vienna, Austria, September 23-27 2007.

[151] M. Leman. Schema-based tone center recognition of musical signals. *Journal of New Music Research*, 23(2):169–204, 1994.

[152] Marc Leman, Lieven Clarisse, Bernard De Baets, Hans De Meyer, Micheline Lesaffre, Gaë tan Martens, Jean Martens, and D Van Steelant. Tendencies, perspectives, and opportunities of musical audio-mining. In A Calvo-Manzano, A Perez-Lopez, and J Salvador Santiago, editors, *REVISTA DE ACUSTICA*, volume 33, pages [1]–[6], 2002.

[153] Mark Levy and Mark Sandler. A semantic space for music derived from social tags. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, Vienna, Austria, September 2007.

[154] Michael S. Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain. Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 2(1):1–19, February 2006.

[155] Cynthia CS Liem, Nicola Orio, Geoffroy Peeters, and Markus Schedl. Brave new task: Musiclef multimodal music tagging. In *Working Notes Proceedings of the MediaEval 2012 Workshop*, Pisa, Italy, 2012.

[156] G. Linden, B. Smith, and J. York. Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing*, 4(1), 2003.

[157] S. Lippens, J.P. Martens, M. Leman, B. Baets, H. Mayer, and G. Tzanetakis. A comparison of human and automatic musical genre classification. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pages iv–233–iv–236 vol.4, 2004.

[158] Beth Logan. Mel Frequency Cepstral Coefficients for Music Modeling. In *Proceedings of the International Symposium on Music Information Retrieval (ISMIR 2000)*, Plymouth, MA, USA, October 2000.

[159] Beth Logan, Daniel P.W. Ellis, and Adam Berenzweig. Toward Evaluation Techniques for Music Similarity. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2003): Workshop on the Evaluation of Music Information Retrieval Systems*, Toronto, Canada, July–August 2003. ACM Press.

[160] David G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the International Conference on Computer Vision (ICCV) 2*, ICCV '99, pages 1150–1157, Washington, DC, USA, 1999. IEEE Computer Society.

[161] Dominik Lübbers and Matthias Jarke. Adaptive Multimodal Exploration of Music Collections. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, Kobe, Japan, October 2009.

[162] Robert C Maher and James W Beauchamp. Fundamental frequency estimation of musical signals using a two-way mismatch procedure. *The Journal of the Acoustical Society of America*, 95:2254, 1994.

[163] Michael I. Mandel, Douglas Eck, and Yoshua Bengio. Learning tags that vary within a song. In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, pages 399–404, Utrecht, the Netherlands, August 2010.

[164] Michael I. Mandel and Daniel P. W. Ellis. A Web-Based Game for Collecting Music Metadata. *Journal of New Music Research*, 37(2):151–165, 2008.

[165] Michael I. Mandel, Razvan Pascanu, Douglas Eck, Yoshua Bengio, Luca M. Aiello, Rossano Schifanella, and Filippo Menczer. Contextual Tag Inference. *ACM Transactions on Multimedia Computing, Communications and Applications*, 7S(1):32:1–32:18, 2011.

[166] Arpi Mardirossian and Elaine Chew. Visualizing Music: Tonal Progressions and Distributions. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, pages 189–194, Vienna, Austria, September 23-27 2007.

[167] A. Martorell and Emilia Gómez. Two-dimensional visual inspection of pitch-space, many time-scales and tonal uncertainty over time. In *3rd International Conference on Mathematics and Computation in Music (MCM 2011)*, Paris, 15/06/2011 2011.

[168] Matthias Mauch. *Automatic chord transcription from audio using computational models of musical context*. PhD thesis, School of Electronic Engineering and Computer Science Queen Mary, University of London, 2010.

[169] Rudolf Mayer and Andreas Rauber. Towards time-resilient mir processes. In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, pages 337–342, Porto, Portugal, October 2012.

[170] Brian McFee, Thierry Bertin-Mahieux, Dan Ellis, and Gert Lanckriet. The million song dataset challenge. In *Proc. of the 4th International Workshop on Advances in Music Information Research (AdMIRe)*, April 2012.

[171] Martin F McKinney and Dirk Moelants. Extracting the perceptual tempo from music. In *Proc. Int. Conf. on Music Info. Retr. ISMIR-04*, pages 146–149, 2004.

[172] Riccardo Miotto, Luke Barrington, and Gert Lanckriet. Improving Auto-tagging by Modeling Semantic Co-occurrences. In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, Utrecht, the Netherlands, August 2010.

[173] Riccardo Miotto and Nicola Orio. A probabilistic model to combine tags and acoustic similarity for music retrieval. *ACM Transactions on Information Systems*, 30(2):8:1–8:29, May 2012.

[174] Dirk Moelants and M McKinney. Tempo perception and musical content: What makes a piece fast, slow or temporally ambiguous. In *Proceedings of the 8th International Conference on Music Perception and Cognition*, pages 558–562, 2004.

[175] Bart Moens, Leon van Noorden, and Marc Leman. D-Jogger: Syncing Music with Walking. In *Proceedings of the 7th Sound and Music Computing Conference (SMC)*, pages 451–456, Barcelona, Spain, 2010.

[176] Henning Müller, Paul Clough, Thomas Deselaers, Barbara Caputo, and Image CLEF. Imageclef: Experimental evaluation in visual information retrieval. *The Information Retrieval Series*, 32, 2010.

[177] Meinard Müller and Sebastian Ewert. Towards timbre-invariant audio features for harmony-based music. *Audio, Speech, and Language Processing, IEEE Transactions on*, 18(3):649–662, 2010.

[178] Meinard Müller, Masataka Goto, and Markus Schedl, editors. *Multimodal Music Processing*, volume 3 of *Dagstuhl Follow-Ups*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2012.

[179] Meinard Müller and Nanzhu Jiang. A scape plot representation for visualizing repetitive structures of music recordings. In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, Porto, Portugal, October 8-12 2012.

[180] Meinard Müller, Henning Mattes, and Frank Kurth. An efficient multiscale approach to audio synchronization. In *International Conference on Music Information Retrieval*, pages 192–197, Victoria, Canada, 2006.

[181] Bernhard Niedermayer. *Accurate Audio-to-Score Alignment — Data Acquisition in the Context of Computational Musicology*. PhD thesis, Johannes Kepler University, Linz, Austria, 2012.

[182] Bernhard Niedermayer, Sebastian Böck, and Gerhard Widmer. On the importance of "real" audio data for mir algorithm evaluation at the note level: a comparative study. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, Miami, Florida, USA, October 2011.

[183] A Michael Noll. Cepstrum pitch determination. *The journal of the acoustical society of America*, 41:293, 1967.

[184] Andreas Nürnberger and Marcin Detyniecki. Weighted Self-Organizing Maps: Incorporating User Feedback. In Okyay Kaynak and Erkki Oja, editors, *Proceedings of the Joined 13th International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP 2003)*, pages 883–890. Springer-Verlag, 2003.

[185] Bee Suan Ong et al. Structural analysis and segmentation of music signals. 2007.

[186] Nicola Orio. Music retrieval: A tutorial and review. *Foundations and Trends in Information Retrieval*, 1(1):1–90, 2006.

[187] Nicola Orio, David Rizo, Riccardo Miotto, Nicola Montecchio, Markus Schedl, and Olivier Lartillot. Musiclef: A benchmark activity in multimodal music information retrieval. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, Miami, Florida, USA, October 2011.

[188] Laurent Oudre, Yves Grenier, and Cédric Févotte. Template-based chord recognition: Influence of the chord types. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, pages 153–158, 2009.

[189] Francois Pachet and Jean-Julien Aucouturier. Improving timbre similarity: How high is the sky? *Journal of negative results in speech and audio sciences*, 1(1):1–13, 2004.

[190] François Pachet and Pierre Roy. Hit Song Science is Not Yet a Science. In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR 2008)*, Philadelphia, PA, USA, September 2008.

[191] François Pachet, Gert Westerman, and Damien Laigre. Musical Data Mining for Electronic Music Distribution. In *Proceedings of the 1st International Conference on Web Delivering of Music (WEDELMUSIC 2001)*, Florence, Italy, November 23–24 2001.

[192] Kevin R. Page, Ben Fields, David de Roure, Tim Crawford, and J. Stephen Downie. Capturing the workflows of music information retrieval for repeatability and reuse. *International Journal on Intelligent Information Systems*, 2013.

[193] Elias Pampalk. Islands of Music: Analysis, Organization, and Visualization of Music Archives. Master's thesis, Vienna University of Technology, Vienna, Austria, 2001. http://www.oefai.at/∼elias/music/thesis.html.

[194] Elias Pampalk. Audio-based music similarity and retrieval: Combining a spectral similarity model with information extracted from fluctuation patterns. In *Proceedings of the International Symposium on Music Information Retrieval*, 2006.

[195] Elias Pampalk and Masataka Goto. MusicRainbow: A New User Interface to Discover Artists Using Audio-based Similarity and Web-based Labeling. In *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR 2006)*, Victoria, Canada, October 8–12 2006.

[196] Elias Pampalk, Tim Pohle, and Gerhard Widmer. Dynamic Playlist Generation Based on Skipping Behavior. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR)*, London, UK, September 2005.

[197] Hélene Papadopoulos and Geoffroy Peeters. Large-scale study of chord estimation algorithms based on chroma representation and hmm. In *Content-Based Multimedia Indexing, 2007. CBMI'07. International Workshop on*, pages 53–60. IEEE, 2007.

[198] Hélène Papadopoulos, Geoffroy Peeters, et al. Local key estimation based on harmonic and metric structures. In *Proceedings of the International Conference on Digital Audio Effects*, pages 408–415, 2009.

[199] Johan Pauwels and Geoffroy Peeters. Evaluating automatically estimated chord sequences. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 749–753, 2013.

[200] G. Peeters, B.L. Giordano, P. Susini, N. Misdariis, and S. McAdams. The timbre toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America*, 130(5):2902?916, 2011.

[201] Geoffroy Peeters. A large set of audio features for sound description (similarity and classification) in the cuidado project. *CUIDADO internal report*, 2004.

[202] Geoffroy Peeters, Amaury La Burthe, and Xavier Rodet. Toward automatic music audio summary generation from signal analysis. In *In Proc. International Conference on Music Information Retrieval*, pages 94–100, 2002.

[203] Geoffroy Peeters and Karen Fort. Towards a (Better) Definition of the Description of Annotated MIR Corpora. In *International Society for Music Information Retrieval Conference*, pages 25–30, 2012.

[204] Geoffroy Peeters, Julián Urbano, and Gareth J.F. Jones. Notes from the ismir 2012 late-breaking session on evaluation in music information retrieval. In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, Porto, Portugal, October 2012.

[205] T. Pohle, P. Knees, M. Schedl, and G. Widmer. Building an Interactive Next-Generation Artist Recommender Based on Automatically Derived High-Level Concepts. In *Proc. CBMI*, 2007.

[206] Tim Pohle, Peter Knees, Markus Schedl, Elias Pampalk, and Gerhard Widmer. “Reinventing the Wheel”: A Novel Approach to Music Player Interfaces. *IEEE Transactions on Multimedia*, 9:567–575, 2007.

[207] Graham E Poliner, Daniel PW Ellis, Andreas F Ehmann, Emilia Gómez, Sebastian Streich, and Beesuan Ong. Melody transcription from music audio: Approaches and evaluation. *Audio, Speech, and Language Processing, IEEE Transactions on*, 15(4):1247–1256, 2007.

[208] L.R. Rabiner and R.W. Schafer. *Introduction to digital speech processing*. Foundations and Trends in Signal Processing Series. Now the essence of knowledge, 2007.

[209] Edie Rasmussen. Evaluation in Information Retrieval. In *Panel on Music Information Retrieval Evaluation Frameworks at ISMIR 2002*, pages 43–44, 2002.

[210] Andreas Rauber, Alexander Schindler, and Rudolf Mayer. Facilitating comprehensive benchmarking experiments on the million song dataset. In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, pages 469–474, Porto, Portugal, October 2012.

[211] Gordon Reynolds, Dan Barry, Ted Burke, and Eugene Coyle. Towards a Personal Automatic Music Playlist Generation Algorithm: The Need for Contextual Information. In *Proceedings of the 2nd International Audio Mostly Conference: Interaction with Sound*, pages 84–89, Ilmenau, Germany, 2007.

[212] Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor, editors. *Recommender Systems Handbook*. Springer, 2011.

[213] Matei Ripeanu. Peer-to-Peer Architecture Case Study: Gnutella Network. In *Proceedings of the IEEE International Conference on Peer-to-Peer Computing (P2P 2001)*, Linköping, Sweden, August 2001. IEEE.

[214] Joseph J. Rocchio. Relevance Feedback in Information Retrieval. In Gerard Salton, editor, *The SMART Retrieval System - Experiments in Automatic Document Processing*, pages 313–323. Englewood Cliffs, NJ: Prentice-Hall, 1971.

[215] J. Salamon and E. Gómez. Melody extraction from polyphonic music signals using pitch contour characteristics. *IEEE Transactions on Audio, Speech and Language Processing*, 20:1759–1770, 08/2012 2012.

[216] J. Salamon, Emilia Gómez, D. P. W. Ellis, and G. Richard. Melody extraction from polyphonic music signals: Approaches, applications and challenges. *IEEE Signal Processing Magazine*, In Press.

[217] J. Salamon, Sankalp Gulati, and Xavier Serra. A multipitch approach to tonic identification in indian classical music. In *13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, pages 499–504, Porto, 08/10/2012 2012.

[218] Justin Salamon, Joan Serra, and Emilia Gómez. Tonal representations for music retrieval: from version identification to query-by-humming. *International Journal of Multimedia Information Retrieval*, pages 1–14, 2013.

[219] Justin Salamon and Julián Urbano. Current challenges in the evaluation of predominant melody extraction algorithms. In *International Society for Music Information Retrieval Conference*, pages 289–294, Porto, Portugal, October 2012.

[220] Phillipe Salembier, Thomas Sikora, and BS Manjunath. *Introduction to MPEG-7: multimedia content description interface*. John Wiley & Sons, Inc., 2002.

[221] Gerard Salton, A. Wong, and C. S. Yang. A Vector Space Model for Automatic Indexing. *Communications of the ACM*, 18(11):613–620, 1975.

[222] Mark Sanderson. Test collection based evaluation of information retrieval systems. *Foundations and Trends in Information Retrieval*, 4(4):247–375, 2010.

[223] Craig Stuart Sapp. Visual hierarchical key analysis. *Computers in Entertainment (CIE)*, 3(4):1–19, 2005.

[224] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl. Item-based Collaborative Filtering Recommendation Algorithms. In *Proc. WWW*, 2001.

[225] Nicolas Scaringella, Giorgio Zoia, and Daniel Mlynek. Automatic genre classification of music content: a survey. *Signal Processing Magazine, IEEE*, 23(2):133–141, 2006.

[226] Markus Schedl. *Automatically Extracting, Analyzing, and Visualizing Information on Music Artists from the World Wide Web*. PhD thesis, Johannes Kepler University Linz, Linz, Austria, 2008.

[227] Markus Schedl. *Music Data Mining*, chapter Web-Based and Community-Based Music Information Extraction. CRC Press/Chapman Hall, 2011.

[228] Markus Schedl. #nowplaying Madonna: A Large-Scale Evaluation on Estimating Similarities Between Music Artists and Between Movies from Microblogs. *Information Retrieval*, 15:183–217, June 2012.

[229] Markus Schedl. Leveraging Microblogs for Spatiotemporal Music Information Retrieval. In *Proceedings of the 35th European Conference on Information Retrieval (ECIR 2013)*, Moscow, Russia, March 24–27 2013.

[230] Markus Schedl, Arthur Flexer, and Julián Urbano. The Neglected User in Music Information Retrieval Research. *International Journal of Journal of Intelligent Information Systems*, 2013.

[231] Markus Schedl, David Hauger, and Dominik Schnitzer. A Model for Serendipitous Music Retrieval. In *Proceedings of the 16th International Conference on Intelligent User Interfaces (IUI 2012): 2nd International Workshop on Context-awareness in Retrieval and Recommendation (CaRR 2012)*, Lisbon, Portugal, February 14 2012.

[232] Markus Schedl, David Hauger, and Julián Urbano. Harvesting microblogs for contextual music similarity estimation - a co-occurrence-based framework. *Multimedia Systems*, 2013.

[233] Markus Schedl, Peter Knees, Tim Pohle, and Gerhard Widmer. Towards Automatic Retrieval of Album Covers. In *Proceedings of the 28th European Conference on Information Retrieval (ECIR 2006)*, London, UK, April 2–5 2006.

[234] Markus Schedl, Peter Knees, and Gerhard Widmer. A Web-Based Approach to Assessing Artist Similarity using Co-Occurrences. In *Proceedings of the 4th International Workshop on Content-Based Multimedia Indexing (CBMI 2005)*, Riga, Latvia, June 21–23 2005.

[235] Markus Schedl, Cynthia C.S. Liem, Geoffroy Peeters, and Nicola Orio. A Professionally Annotated and Enriched Multimodal Data Set on Popular Music. In *Proceedings of the 4th ACM Multimedia Systems Conference (MMSys 2013)*, Oslo, Norway, February–March 2013.

[236] Markus Schedl, Tim Pohle, Peter Knees, and Gerhard Widmer. Exploring the Music Similarity Space on the Web. *ACM Transactions on Information Systems*, 29(3), July 2011.

[237] Markus Schedl, Tim Pohle, Noam Koenigstein, and Peter Knees. What's Hot? Estimating Country-Specific Artist Popularity. In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, Utrecht, the Netherlands, August 2010.

[238] Markus Schedl and Dominik Schnitzer. Hybrid Retrieval Approaches to Geospatial Music Recommendation. In *Proceedings of the 35th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, Dublin, Ireland, July 31–August 1 2013.

[239] Markus Schedl and Dominik Schnitzer. Location-Aware Music Artist Recommendation. In *Proceedings of the 20th International Conference on MultiMedia Modeling (MMM 2014)*, Dublin, Ireland, January 2014.

[240] Markus Schedl, Klaus Seyerlehner, Dominik Schnitzer, Gerhard Widmer, and Cornelia Schiketanz. Three Web-based Heuristics to Determine a Person's or Institution's Country of Origin. In *Proceedings of the 33th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR2010)*, Geneva, Switzerland, July 19–23 2010.

[241] Markus Schedl, Sebastian Stober, Emilia Gómez, Nicola Orio, and Cynthia C.S. Liem. User-Aware Music Retrieval. In Meinard Müller, Masataka Goto, and Markus Schedl, editors, *Multimodal Music Processing*, volume 3 of *Dagstuhl Follow-Ups*, pages 135–156. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2012.

[242] Markus Schedl and Gerhard Widmer. Automatically Detecting Members and Instrumentation of Music Bands via Web Content Mining. In *Proceedings of the 5th Workshop on Adaptive Multimedia Retrieval (AMR 2007)*, Paris, France, July 5–6 2007.

[243] Markus Schedl, Gerhard Widmer, Peter Knees, and Tim Pohle. A music information system automatically generated via web content mining techniques. *Information Processing & Management*, 47, 2011.

[244] Mark A Schmuckler. Pitch and pitch structures. *Ecological psychoacoustics*, pages 271–315, 2004.

[245] Dominik Schnitzer, Arthur Flexer, Markus Schedl, and Gerhard Widmer. Local and Global Scaling Reduce Hubs in Space. *Journal of Machine Learning Research*, 13:2871–2902, October 2012.

[246] Dominik Schnitzer, Tim Pohle, Peter Knees, and Gerhard Widmer. One-Touch Access to Music on Mobile Devices. In *Proceedings of the 6th International Conference on Mobile and Ubiquitous Multimedia (MUM 2007)*, Oulu, Finland, December 12–14 2007.

[247] Eleanor Selfridge-Field. Conceptual and Representational Issues in Melodic Comparison. *Computing in Musicology*, 11:3–64, 1998.

[248] J. Serrà, E. Gómez, and P. Herrera. *Audio cover song identification and similarity: background, approaches, evaluation, and beyond*, volume 274 of *Studies in Computational Intelligence*, chapter 14, pages 307–332. Springer-Verlag Berlin / Heidelberg, 2010. Please note that the PDF linked here is a preliminary draft. You can access the final revised version through Springerlink: http://www.springerlink.com/content/a02r21125nw63551/.

[249] Xavier Serra. Data Gathering for a Culture Specific Approach in MIR. In *Proceedings of the 21st International World Wide Web Conference (WWW 2012): 4th International Workshop on Advances in Music Information Research (AdMIRe 2012)*, Lyon, France, April 17 2012.

[250] Xavier Serra, Michela Magas, Emmanouil Benetos, Magdalena Chudy, S. Dixon, Arthur Flexer, Emilia Gómez, F. Gouyon, P. Herrera, S. Jordà, Oscar Paytuvi, G. Peeters, Jan Schlüter, H. Vinet, and G. Widmer. *Roadmap for Music Information ReSearch*. 2013.

[251] William A Sethares. Local consonance and the relationship between timbre and scale. *The Journal of the Acoustical Society of America*, 94:1218, 1993.

[252] Klaus Seyerlehner. *Content-Based Music Recommender Systems: Beyond simple Frame-Level Audio Similarity*. PhD thesis, Johannes Kepler University Linz, Linz, Austria, 2010.

[253] Klaus Seyerlehner, Markus Schedl, Peter Knees, and Reinhard Sonnleitner. A Refined Block-Level Feature Set for Classification, Similarity and Tag Prediction. In *Extended Abstract to the Music Information Retrieval Evaluation eXchange (MIREX 2011) / 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, Miami, FL, USA, October 2009.

[254] Klaus Seyerlehner, Gerhard Widmer, and Tim Pohle. Fusing Block-Level Features for Music Similarity Estimation. In *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx-10)*, Graz, Austria, September 2010.

[255] Yuval Shavitt and Udi Weinsberg. Songs Clustering Using Peer-to-Peer Co-occurrences. In *Proceedings of the IEEE International Symposium on Multimedia (ISM2009): International Workshop on Advances in Music Information Research (AdMIRe 2009)*, San Diego, CA, USA, December 16 2009.

[256] S. Sigurdsson, K. B Petersen, and T. Lehn-Schiøler. Mel frequency cepstral coefficients: An evaluation of robustness of mp3 encoded music. In *International Conference on Music Information Retrieval (ISMIR'07)*, pages 286–289, 2006.

[257] Carlos N Silla Jr, Alessandro L Koerich, and Celso AA Kaestner. The latin music database. In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR 2008)*, pages 451–456, Philadelphia, PA, USA, September 2008.

[258] Janto Skowronek, Martin McKinney, and Steven Van De Par. Groundtruth for automatic music mood classification. In *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR 2006)*, pages 395–396, Victoria, Canada, October 8–12 2006.

[259] Malcolm Slaney. Web-Scale Multimedia Analysis: Does Content Matter? *IEEE MultiMedia*, 18(2):12–15, 2011.

[260] Alan F. Smeaton, Paul Over, and Wessel Kraaij. Evaluation campaigns and trecvid. In *MIR '06: Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval*, pages 321–330, New York, NY, USA, 2006. ACM Press.

[261] Jordan BL Smith and Elaine Chew. A meta-analysis of the mirex structure segmentation task. In *Proc. of the 14th International Society for Music Information Retrieval Conference, Curitiba, Brazil*, 2013.

[262] Paul Smolensky. *Information Processing in Dynamical Systems: Foundations of Harmony Theory*, pages 194–281. MIT Press, Cambridge, MA, USA, 1986.

[263] Mohammad Soleymani, Michael N. Caro, Erik M. Schmidt, and Yi-Hsuan Yang. The mediaeval 2013 brave new task: Emotion in music. In *MediaEval*, 2013.

[264] Mohamed Sordo. *Semantic Annotation of Music Collections: A Computational Approach.* PhD thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2012.

[265] Mohamed Sordo, Oscar Celma, Martın Blech, and Enric Guaus. The quest for musical genres: do experts and the wisdom of crowds agree? In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR 2008)*, pages 255–260, Philadelphia, PA, USA, September 2008.

[266] Mohamed Sordo, Òscar Celma, and Cyril Laurier. QueryBag: Using Different Sources For Querying Large Music Collections. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR)*, Kobe, Japan, October 2009.

[267] Adam M Stark, Matthew EP Davies, and Mark D Plumbley. Real-time beatsynchronous analysis of musical audio. In *Proceedings of the 12th Int. Conference on Digital Audio Effects, Como, Italy*, pages 299–304, 2009.

[268] Rebecca Stewart and Mark Sandler. The amblr: A Mobile Spatial Audio Music Browser. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, Barcelona, Spain, 2011.

[269] Sebastian Stober. *Adaptive Methods for User-Centered Organization of Music Collections*. PhD thesis, Otto-von-Guericke-University, Magdeburg, Germany, November 2011. published by Dr. Hut Verlag, ISBN 978-3-8439-0229-8.

[270] Sebastian Stober and Andreas Nürnberger. MusicGalaxy: A Multi-focus Zoomable Interface for Multi-facet Exploration of Music Collections. In *Proceedings of the 7th International Symposium on Computer Music Modeling and Retrieval (CMMR 2010)*, Málaga, Spain, June 21–24 2010.

[271] Bob L. Sturm. Two systems for automatic music genre recognition. In *Proceedings of the second international ACM workshop on Music information retrieval with user-centered and multimodal strategies*, pages 69–74, 2012.

[272] Bob L. Sturm. Classification accuracy is not enough. *Journal of Intelligent Information Systems*, 2013.

[273] Jean Tague-Sutcliffe. The Pragmatics of Information Retrieval Experimentation, Revisited. *Information Processing and Management*, 28(4):467–490, 1992.

[274] David Temperley. What's key for key? the krumhansl-schmuckler key-finding algorithm reconsidered. *Music Perception*, pages 65–100, 1999.

[275] David Temperley. A bayesian key-finding model. In *2005 MIREX Contest - Symbolic Key Finding*, 2005. http://www.music-ir.org/evaluation/mirex-results/sym-key/index.html.

[276] Yuan-Ching Teng, Ying-Shu Kuo, and Yi-Hsuan Yang. A large in-situ dataset for context-aware music recommendation on smartphones. In *IEEE International Conference on Multimedia and Expo Workshops (ICMEW) 2013*, pages 1–4, San Jose, CA, USA, July 2013.

[277] D. Turnbull, R. Liu, L. Barrington, and G. Lanckriet. A Game-based Approach for Collecting Semantic Annotations of Music. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, Vienna, Austria, September 2007.

[278] Douglas Turnbull, Luke Barrington, Mehrdad Yazdani, and Gert Lanckriet. Combining Audio Content and Social Context for Semantic Music Discovery. In *Proceedings of the 32th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, Boston, MA, USA, July 2009.

[279] Rainer Typke, Marc den Hoed, Justin de Nooijer, Frans Wiering, and Remco C. Veltkamp. A ground truth for half a million musical incipits. *Journal of Digital Information Management*, 3(1):34–39, 2005.

[280] Rainer Typke, Remco C. Veltkamp, and Frans Wiering. A measure for evaluating retrieval techniques based on partially ordered ground truth lists. In *IEEE International Conference on Multimedia and Expo*, pages 1793–1796, 2006.

[281] George Tzanetakis and Perry Cook. Musical Genre Classification of Audio Signals. *IEEE Transactions on Speech and Audio Processing*, 10(5):293–302, 2002.

[282] Julián Urbano. *Evaluation in Audio Music Similarity.* PhD thesis, University Carlos III of Madrid, 2013.

[283] Julián Urbano, Dmitry Bogdanov, Perfecto Herrera, Emilia Gómez, and Xavier Serra. What is the effect of audio quality on the robustness of MFCCs and chroma features? In *International Society for Music Information Retrieval Conference*, Taipei, Taiwan, October 2014.

[284] Julián Urbano, J. Stephen Downie, Brian Mcfee, and Markus Schedl. How significant is statistically significant? the case of audio music similarity and retrieval. In *International Society for Music Information Retrieval Conference*, pages 181–186, Porto, Portugal, October 2012.

[285] Julián Urbano, Mónica Marrero, Diego Martín, and Juan Lloréns. Improving the generation of ground truths based on partially ordered lists. In *International Society for Music Information Retrieval Conference*, pages 285–290, Utrecht, the Netherlands, August 2010.

[286] Julián Urbano, Diego Martín, Mónica Marrero, and Jorge Morato. Audio music similarity and retrieval: Evaluation power and stability. In *International Society for Music Information Retrieval Conference*, pages 597–602, Miami, Florida, USA, October 2011.

[287] Julián Urbano, Jorge Morato, Mónica Marrero, and Diego Martín. Crowdsourcing preference judgments for evaluation of music similarity tasks. In *Proceedings of the 1st ACM SIGIR Workshop on Crowdsourcing for Search Evaluation*, page 9?6, 2010.

[288] Julián Urbano and Markus Schedl. Minimal Test Collections for Low-Cost Evaluation of Audio Music Similarity and Retrieval Systems. *International Journal of Multimedia Information Retrieval*, 2(1):59–70, 2013.

[289] Julián Urbano, Markus Schedl, and Xavier Serra. Evaluation in music information retrieval. *International Journal on Intelligent Information Systems*, 2013.

[290] Vladimir N. Vapnik. *Statistical Learning Theory.* Wiley, Chichester, UK, 1998.

[291] Nuno Vasconcelos. Image Indexing with Mixture Hierarchies. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Kauai, Hawaii, June 2001.

[292] Fabio Vignoli and Steffen Pauws. A music retrieval system based on user driven similarity and its evaluation. In *ISMIR*, pages 272–279. Citeseer, 2005.

[293] Luis von Ahn and Laura Dabbish. Labeling Images with a Computer Game. In *CHI'04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 319–326, New York, NY, USA, 2004. ACM Press.

[294] Ellen M. Voorhees. *Whither Music IR Evaluation Infrastructure: Lessons to be Learned from TREC*, pages 7–3. 2002.

[295] Ellen M. Voorhees and Donna K. Harman. *TREC: Experiment and Evaluation in Information Retrieval.* MIT Press, 2005.

[296] Gregory H. Wakefield. Mathematical representation of joint time-chroma distributions. In *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, pages 637–645, 1999.

[297] Avery Li-Chun Wang. An Industrial Strength Audio Search Algorithm. In *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2003)*, Baltimore, Maryland, USA, October 26–30 2003.

[298] Xinxi Wang, David Rosenblum, and Ye Wang. Context-aware mobile music recommendation for daily activities. In *Proceedings of the 20th ACM International Conference on Multimedia*, pages 99–108, New York, NY, USA, 2012. ACM.

[299] Christian Wartena. Comparing segmentation strategies for efficient video passage retrieval. In *Content-Based Multimedia Indexing (CBMI), 2012 10th International Workshop on*, pages 1–6. IEEE, 2012.

[300] David M Weigl and Catherine Guastavino. User studies in the music information retrieval literature. In *Proceedings of the 12th International Society for Music Information Retrieval conference (ISMIR 2011), Miami, USA*, 2011.

[301] David L Wessel. Timbre space as a musical control structure. *Computer music journal*, 3(2):45–52, 1979.

[302] Brian Whitman and Steve Lawrence. Inferring Descriptions and Similarity for Music from Community Metadata. In *Proceedings of the 2002 International Computer Music Conference (ICMC 2002)*, pages 591–598, Göteborg, Sweden, September 16–21 2002.

[303] Daniel Wolff and Tillman Weyde. Adapting Metrics for Music Similarity using Comparative Ratings. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, Miami, FL, USA, October 2011.

[304] Yi-Hsuan Yang and Homer H. Chen. *Music Emotion Recognition*. CRC Press, 2011.

[305] Yi-Hsuan Yang and Homer H. Chen. Machine recognition of music emotion: A review. *Transactions on Intelligent Systems and Technology*, 3(3), May 2013.

[306] Chunghsin Yeh, Axel Roebel, and Xavier Rodet. Multiple fundamental frequency estimation and polyphony inference of polyphonic music signals. *Audio, Speech, and Language Processing, IEEE Transactions on*, 18(6):1116–1126, 2010.

[307] Yuan Cao Zhang, Diarmuid O Seaghdha, Daniele Quercia, Tamas Jambor. Auralist: Introducing Serendipity into Music Recommendation. In *Proceedings of the 5th ACM Int'l Conference on Web Search and Data Mining (WSDM)*, Seattle, WA, USA, February 8–12 2012.

[308] Mark Zadel and Ichiro Fujinaga. Web Services for Music Information Retrieval. In *Proceedings of the 5th International Symposium on Music Information Retrieval (ISMIR 2004)*, Barcelona, Spain, October 10–14 2004.

[309] Eva Zangerle, Wolfgang Gassler, and Günther Specht. Exploiting Twitter's Collective Knowledge for Music Recommendations. In *Proceedings of the 21st International World Wide Web Conference (WWW 2012): Making Sense of Microposts (#MSM2012)*, pages 14–17, Lyon, France, April 17 2012.

[310] Bingjun Zhang, Jialie Shen, Qiaoliang Xiang, and Ye Wang. CompositeMap: A Novel Framework for Music Similarity Measure. In *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2009)*, pages 403–410, New York, NY, USA, 2009. ACM.

[311] Justin Zobel, William Webber, Mark Sanderson, and Alistair Moffat. Principles for Robust Evaluation Infrastructure. In *ACM CIKM Workshop on Data infrastructures for Supporting Information Retrieval Evaluation*, 2011.