Full text available at: http://dx.doi.org/10.1561/1500000002

Music Retrieval: A Tutorial and Review

Full text available at: http://dx.doi.org/10.1561/1500000002

Music Retrieval: A Tutorial and Review

Nicola Orio

Department of Information Engineering
University of Padova
Via Gradenigo, 6/b, Padova 35131, Italy
orio@dei.unipd.it

now
the essence of knowledge
Boston – Delft

Full text available at: http://dx.doi.org/10.1561/1500000002

Foundations and Trends® in Information Retrieval

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
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

Library of Congress Control Number: 2006935708

The preferred citation for this publication is Nicola Orio, Music Retrieval: A Tutorial and Review, Foundation and Trends® in Information Retrieval, vol 1, no 1, pp 1–96, November 2006

Printed on acid-free paper

ISBN: 1-933019-39-5
© November 2006 Nicola Orio

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, mechanical, photocopying, recording or otherwise, without prior written permission of the publishers.

Photocopying. In the USA: This journal is registered at the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923. Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by now Publishers Inc for users registered with the Copyright Clearance Center (CCC). The 'services' for users can be found on the internet at: www.copyright.com

For those organizations that have been granted a photocopy license, a separate system of payment has been arranged. Authorization does not extend to other kinds of copying, such as that for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. In the rest of the world: Permission to photocopy must be obtained from the copyright owner. Please apply to now Publishers Inc., PO Box 1024, Hanover, MA 02339, USA; Tel. +1 781 871 0245; www.nowpublishers.com; sales@nowpublishers.com

now Publishers Inc. has an exclusive license to publish this material worldwide. Permission to use this content must be obtained from the copyright license holder. Please apply to now Publishers, PO Box 179, 2600 AD Delft, The Netherlands, www.nowpublishers.com; e-mail: sales@nowpublishers.com

Full text available at: http://dx.doi.org/10.1561/1500000002

Foundations and Trends® in Information Retrieval
Volume 1 Issue 1, November 2006

Editorial Board

Editors-in-Chief:
Fabrizio Sebastiani
Consiglio Nazionale delle Ricerche
fabrizio.sebastiani@isti.cnr.it

Jamie Callan
Carnegie Mellon University
callan@cmu.edu

Editors
Alan Smeaton (Dublin City University)
Andrei Z. Broder (Yahoo! Research)
Bruce Croft (University of Massachusetts, Amherst)
Charles L.A. Clarke (University of Waterloo)
Ellen Voorhees (National Institute of Standards and Technology)
Ian Ruthven (University of Strathclyde, Glasgow)
James Allan (University of Massachusetts, Amherst)
Justin Zobel (RMIT University, Melbourne)
Karen Sparck Jones (University of Cambridge)
Maarten de Rijke (University of Amsterdam)
Marcello Federico (ITC-irst)
Norbert Fuhr (University of Duisburg-Essen)
Soumen Chakrabarti (Indian Institute of Technology)
Susan Dumais (Microsoft Research)
Wei-Ying Ma (Microsoft Research Asia)
William W. Cohen (CMU)

Full text available at: http://dx.doi.org/10.1561/1500000002

Editorial Scope

Foundations and Trends® in Information Retrieval will publish 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, November 2006, Volume 1, 4 issues. ISSN paper version 1554-0669. ISSN online version 1554-0677. Also available as a combined paper and online subscription.

Full text available at: http://dx.doi.org/10.1561/1500000002

Foundations and Trends® in
Information Retrieval
Vol. 1, No 1 (November 2006) 1–96
© November 2006 Nicola Orio
DOI: 10.1561/1500000002

now
the essence of knowledge

Music Retrieval: A Tutorial and Review

Nicola Orio

Department of Information Engineering, University of Padova
Via Gradenigo, 6/b, Padova 35131, Italy, orio@dei.unipd.it,

Abstract

The increasing availability of music in digital format needs to be matched by the development of tools for music accessing, filtering, classification, and retrieval. The research area of Music Information Retrieval (MIR) covers many of these aspects. The aim of this paper is to present an overview of this vast and new field. A number of issues, which are peculiar to the music language, are described—including forms, formats, and dimensions of music—together with the typologies of users and their information needs. To fulfil these needs a number of approaches are discussed, from direct search to information filtering and clustering of music documents. An overview of the techniques for music processing, which are commonly exploited in many approaches, is also presented. Evaluation and comparisons of the approaches on a common benchmark are other important issues. To this end, a description of the initial efforts and evaluation campaigns for MIR is provided.

Full text available at: http://dx.doi.org/10.1561/1500000002

Contents

Full text available at: http://dx.doi.org/10.1561/1500000002

Full text available at: http://dx.doi.org/10.1561/1500000002

1

Introduction

The amount of available digital music is continuously increasing, promoted by a growing interest of users and by the development of new technology for the ubiquitous enjoyment of music.

There are a number of reasons that may explain this trend, first of all, the characteristics of music language itself. Music is an art form that can be shared by people with different culture because it crosses the barriers of national languages and cultural backgrounds. For example, Western classical music has passionate followers in Japan, and many persons in Europe are keen on classical Indian music: All of them can enjoy music without the need of a translation, which is normally required for accessing foreign textual works. Another reason is that technology for music recording, digitization, and playback allows users for an access that is almost comparable to the listening of a live performance, at least at the level of audio quality, and the signal-to-noise ratio is better for digital formats than for many analog formats. This is not the case with other art forms, like painting or sculpture, for which the digital format is only an approximate representation of the artwork. The access to digitized paintings can be useful for studying the works of a given artist but cannot substitute the direct interaction

Full text available at: http://dx.doi.org/10.1561/1500000002

with the *real world* works. Moreover, music is an art form that can be both cultivated and popular, and sometimes it is impossible to draw a line between the two, as for jazz or for most of traditional music.

Or maybe the increasing interest toward digital music is motivated by its portability and the possibility to access music while doing another, possibly primary, activity. Perhaps users are not looking for a cultural experience, or the enjoyment of artworks, but for a suitable soundtrack for the many hours spent commuting, traveling, waiting or even working, and studying. Last but not least, the music business is pushing toward the continuous production of new musical works, especially in genres like pop and rock. The continuous decrease of the average age of persons that regularly purchase and consume music has been paralleled by an increasing simplification of the structure of mainstream music genres, requiring the continuous production of new music. The number of items sold daily by Web-based music dealers, or downloaded from services like i-Tune—not mentioning the illegal sharing of music files through peer-to-peer networks—shows how much music is commercially and culturally important.

Music Information Retrieval (MIR) is an emerging research area devoted to fulfill users' *music* information needs. As it will be seen, despite the emphasis on retrieval of its name, MIR encompasses a number of different approaches aimed at music management, easy access, and enjoyment. Most of the research work on MIR, of the proposed techniques, and of the developed systems are content based.

The main idea underlying content-based approaches is that a document can be described by a set of features that are directly computed from its content. Usually, content-based access to multimedia data requires specific methodologies that have to be tailored to each particular medium. Yet, the core information retrieval (IR) techniques, which are based on statistics and probability theory, may be more generally employed outside the textual case, because the underlying models are likely to describe fundamental characteristics being shared by different media, languages, and application domains [60]. For this reason, the research results achieved in the area of IR, in particular in the case of text documents, are a continuous reference for MIR approaches. Already in 1996 McLane stated that a challenging research topic would

Full text available at: http://dx.doi.org/10.1561/1500000002

have been the application of some standard principles of text IR to music representation [85].

The basic assumption behind content-based approaches is that metadata are either not suitable, or unreliable, or missing. In the case of MIR, all the three assumptions may hold. As it will be described in more detail in the following sections, music metadata may not be suitable because they are either too generic to discriminate different musical works—there are hundreds of slow ballads sung by a female voice with acoustic guitars in the background—or too specific requiring a precise definition of the information need—name of the singer, name of the album, and date of the first release if not the direct name of the song—or requiring a good musical background—the song is in F lydian, with guitars mainly playing suspended chords, at the end the time signature becomes 3/4 with slower tempo.[1] Metadata can be unreliable in the case of music shared with other users, because there is no control on how the information has been added, and of course metadata may not be present at all.

It is interesting to note that both CD and the first MP3 audio standards, which gave rise to the two digital revolutions of music enjoyment, do not take into account the possibility of carrying also structured metadata information. The possibility to include unstructured textual information in MP3 has been introduced in a later version thanks to an external contribution to the standard. Yet, fields are not mandatory, and it is up to the person who creates the MP3 to spend time adding this information.

Despite these problems with metadata, a number of systems that allow users to access and retrieve music based on textual descriptors have been developed, and are available on the Web. Most of the systems are Digital Libraries devoted to the diffusion of Western classical music, such as Cantate [13], Harmonica [43], and Musica [97], whose names already show the primary interest toward this repertoire—other music digital libraries for Western music are described in [45] and in [27]. The discussion of the relationships between the vast research area of

[1] This sentence has been created on purpose using terms that, although very precise, are not very common outside the music terminology; there is probably no song with these characteristics.

Full text available at: http://dx.doi.org/10.1561/1500000002

digital libraries and MIR would have required a paper by itself, because it regards important issues on digital acquisition of musical works, the creation of an infrastructure for the management of musical documents and for the access to musical content, which are challenging problem by themselves as discussed in [1, 29].

The focus of this paper is on tools, techniques, and approaches for content-based MIR, rather than on systems that implement them. The interested reader can find the descriptions of more than 35 systems for music retrieval in [127], with links to their Web sites. An interesting survey of a selection of 17 different systems has been presented in [135]. Systems can be compared according to the retrieval tasks that can be carried out, the size of the collections, and the techniques that are employed. To this end, a mapping of the analyzed systems on a bidimensional space was proposed [135] in where the two dimensions of interest were the target audience of the systems—e.g., industry, consumer, professional—and the level at which retrieval is needed—e.g., from a particular instance of a work to a music genre.

The paper is structured as follows: This introductory section ends with a short overview of some music concepts. Chapter 2 discusses the peculiarities of the music language, introducing the dimensions that characterize musical documents and that can be used to describe its content. Chapter 3 highlights the main typologies of MIR users, introducing and specifying a number of information needs that have been taken into account by different MIR approaches. The processing of musical documents, aimed at extracting features related to their dimensions of interest, is discussed in Chapter 4, followed by a classification of the different facets of MIR research areas, which are reported in Chapter 5. The efforts carried out for creating a shared evaluation framework and their initial results are presented in Chapter 6. Finally, some concluding considerations are drawn in Chapter 7.

1.1 Review of Music Concepts

Most of the approaches and techniques for music retrieval are based on a number of music concepts that may not be familiar to persons

Full text available at: http://dx.doi.org/10.1561/1500000002

without musical training. For this reason, this section presents a short introduction of some basic concepts and terms.

1.1.1 Basic music elements

Each musical instrument, with the exception of some percussions, produces almost periodic vibrations. In particular, the sounds produced by musical instruments are the result of the combination of different frequencies, which are all multiple integers of a *fundamental frequency*, usually called $F0$.

The three basic features of a musical sound are

- **Pitch**, which is related to the perception of the fundamental frequency of a sound; pitch is said to range from *low* or *deep* to *high* or *acute* sounds.
- **Intensity**, which is related to the amplitude, and thus to the energy, of the vibration; textual labels for intensity range from *soft* to *loud*; the intensity is also defined *loudness*.
- **Timbre**, which is defined as the sound characteristics that allow listeners to perceive as different two sounds with same pitch and same intensity.

The perception of pitch and intensity is not as straightforward as the above definitions may suggest. The human ear does not have a linear behavior neither in pitch recognition nor in the perception of intensity. Yet, these two perceptually relevant qualities of sounds can be reasonably approximated considering the fundamental frequency and the energy of a sound.

Timbre, on the other hand, is a multidimensional sound quality that is not easily described with simple features. Timbre perception is related to the recognition of the sound source—telling a saxophone from a violin—of the playing technique—telling whether a string has been plucked or played with the bow—of the playing technique nuances—the velocity of the bow and its pressure on the string—of the surrounding acoustics—telling whether the violinist has been in a small room or in a concert hall—and of the recording equipment and digital representation—telling whether the sound has been recorded

Full text available at: http://dx.doi.org/10.1561/1500000002

from the broadcast of an AM radio or from a live performance. Given all these characteristics, it is not surprising that timbre has been defined for what it is not.

In the case of many percussive musical instruments, there is no fundamental frequency, and the sound is usually called *noise*. Yet, noises are perceived to be in a low, medium, or high register. Intensity and timbre are still relevant descriptors for noises.

When two or more sounds are played together, they form a *chord*. A chord may have different qualities, depending on the pitch of the different sounds and, in particular, on the distances between them. Chords play a fundamental role in many music genres, in particular in pop, rock, and jazz music, where polyphonic musical instruments—e.g., piano, keyboard, guitar—are often dedicated to the accompaniment and basically play chords.

1.1.2 Music terminology

Apart from the basic concepts introduced in the previous section, there are many terms that are currently used to describe music that may not be familiar to persons without a musical education. Part of the terminology commonly used by the MIR community is derived from music theory and practice. The musical concepts that are relevant for this overview are the following[2]:

- The **tempo** is the speed at which a musical work is played, or expected to be played, by performers. The tempo is usually measured in *beats per minute*.
- The **tonality** of a song is related to the role played by the different chords of a musical work; tonality is defined by the name of the chord that plays a central role in a musical work. The concept of tonality may not be applicable to some music genres.

[2] Being very simple, these operative definitions may not be completely agreed by readers with a musical background, because some relevant aspects are not taken into account. Yet, their purpose is just to introduce some terminology that will be used in the next chapters.

Full text available at: http://dx.doi.org/10.1561/1500000002

- The **time signature**, usually in the form of a fractional number, gives information on the organization of strong and soft beats along the time axis.
- The **key signature**, usually in the form of a number of alterations—symbols ♯ and ♭—is an incomplete representation of the tonality, which is useful for performers because it express which are the notes that have to be consistently played altered.

Figure 1.1 shows four measures of a polyphonic musical score, an excerpt from Claude Debussy's *Première Arabesque*, for piano. In this excerpt no direct information on tempo and tonality is given, the time signature (the **C** sign) indicates that measures have to be divided in four equal beats, and the three sharps (the ♯ signs) indicate that all occurrences of notes *F*, *C*, and *G* have to be raised by a semitone if not otherwise indicated. The presence of three sharps may suggest that the tonality of the excerpt is either A major or F♯ minor, which have the same number of sharps, with the former more likely to be the actual tonality.

Other concepts are more related to the production of sound and to the parameters that describe single or groups of notes. In particular, a note starts being perceived at its *onset time*, lasts for a given *duration*, and stops to be perceived at its *offset time*. Finally, the sounds are produced by musical instruments and by the voice that depending on their conformation have a limited range of pitches that can be produced, which is called instrument—or voice—*register*.

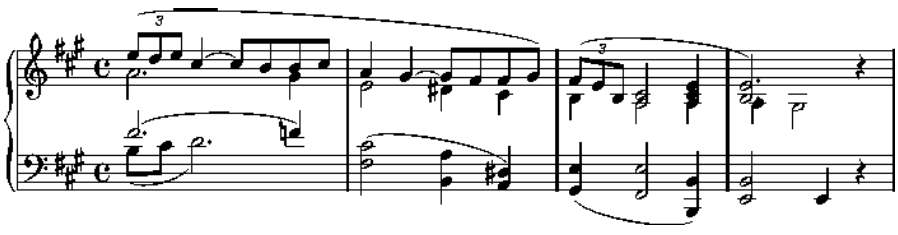

Fig. 1.1 Example of a musical score (excerpt from *Première Arabesque* by Claude Debussy).

Full text available at: http://dx.doi.org/10.1561/1500000002

References

[1] M. Agosti, F. Bombi, M. Melucci, and G.A. Mian. Towards a digital library for the venetian music of the eighteenth century. In J. Anderson, M. Deegan, S. Ross, and S. Harold, editors, *DRH 98: Selected Papers from Digital Resources for the Humanities*, pages 1–16. Office for Humanities Communication, 2000.

[2] E. Allamanche, J. Herre, O. Hellmuth, B. Fröba, T. Kastner, and M. Cremer. Content-based identification of audio material using MPEG-7 low level description. In *Proceedings of the International Symposium on Music Information Retrieval*, pages 73–82, 2001.

[3] P.E. Allen and R.B. Dannenberg. Tracking musical beats in real time. In *Proceedings of the International Computer Music Conference*, pages 140–143, 1990.

[4] D. Bainbridge, C.G. Nevill-Manning, I.H. Witten, L.A. Smith, and R.J. McNab. Towards a digital library of popular music. In *Proceedings of the ACM Conference on Digital Libraries*, pages 161–169, 1999.

[5] M.A. Bartsch and G.H. Wakefield. Audio thumbnailing of popular music using chroma-based representations. *IEEE Transactions on Multimedia*, 7(1):96–104, 2005.

[6] A. Berenzweig, B. Logan, D.P.W. Ellis, and B. Whitman. A large-scale evaluation of acoustic and subjective music-similarity measures. *Computer Music Journal*, 28(2):63–76, 2004.

[7] W.P. Birmingham, R.B. Dannenberg, G.H. Wakefield, M. Bartsch, D. Bykowski, D. Mazzoni, C. Meek, M. Mellody, and W. Rand. MUSART: music retrieval via aural queries. In *Proceedings of the International Conference on Music Information Retrieval*, pages 73–82, 2001.

Full text available at: http://dx.doi.org/10.1561/1500000002

[8] S. Blackburn and D. DeRoure. A tool for content based navigation of music. In *Proceedings of the ACM International Conference on Multimedia*, pages 361–368, 1998.

[9] E. Brazil and M. Fernström. Audio information browsing with the sonic browser. In *Coordinated and Multiple Views in Exploratory Visualization*, pages 26–31, 2003.

[10] E. Cambouropoulos. Musical rhythm: a formal model for determining local boundaries. In E. Leman, editor, *Music, Gestalt and Computing*, pages 277–293. Springer-Verlag, Berlin, DE, 1997.

[11] P. Cano, E. Batlle, T. Kalker, and J. Haitsma. A review of audio fingerprinting. *Journal of VLSI Signal Processing*, 41:271–284, 2005.

[12] P. Cano, A. Loscos, and J. Bonada. Score-performance matching using hmms. In *Proceedings of the International Computer Music Conference*, pages 441–444, 1999.

[13] Cantate. Computer access to notation and text in music libraries, July 2006. http://projects.fnb.nl/cantate/.

[14] M. Clausen, R. Engelbrecht, D. Meyer, and J. Schmitz. PROMS: a web-based tool for searching in polyphonic music. In *Proceedings of the International Symposium of Music Information Retrieval*, 2000.

[15] Coda Music. Enigma transportable file specification. Technical Report, version 98c.0, July 2006. http://www.xs4all.nl/ hanwen/lily-devel/etfspec.pdf.

[16] R.B. Dannenberg and H. Mukaino. New techniques for enhanced quality of computer accompaniment. In *Proceedings of the International Computer Music Conference*, pages 243–249, 1988.

[17] M.E.P. Davies and M.D. Plumbley. Casual tempo tracking of audio. In *Proceedings of the International Conference on Music Information Retrieval*, pages 164–169, 2004.

[18] S.B. Davis and P. Mermelstein. Comparison of parametric representations for monosyllabic word recognition in Continuously Spoken Sentences. *IEEE Transactions on Acoustic, Speech, and Signal Processing*, 28(4):357–366, 1980.

[19] A. de Cheveigné and A. Baskind. F0 estimation. In *Proceedings of Eurospeech*, pages 833–836, 2003.

[20] S. Dixon and G. Widmer. Match: a music alignment tool chest. In *Proceedings of the International Conference of Music Information Retrieval*, pages 492–497, 2005.

[21] S. Doraisamy and S. Rüger. A polyphonic music retrieval system using N-grams. In *Proceedings of the International Conference on Music Information Retrieval*, pages 204–209, 2004.

[22] W.J. Dowling. Scale and contour: two components of a theory of memory for melodies. *Psychological Review*, 85(4):341–354, 1978.

[23] J.S. Downie, J. Futrelle, and D. Tcheng. The International Music Information Retrieval Systems Evaluation Laboratory: governance, access and security. In *Proceedings of the International Conference on Music Information Retrieval*, pages 9–14, 2004.

[24] J.S. Downie, K. West, A. Ehmann, and E. Vincent. The 2005 music information retrieval evaluation exchange (mirex 2005): preliminary overview. In

Full text available at: http://dx.doi.org/10.1561/1500000002

Proceedings of the International Conference on Music Information Retrieval, pages 320–323, 2005.

[25] J.S. Downie. Music information retrieval. *Annual Review of Information Science and Technology*, 37:295–340, 2003.

[26] S. Downie and M. Nelson. Evaluation of a simple and effective music information retrieval method. In *Proceedings of the ACM International Conference on Research and Development in Information Retrieval (SIGIR)*, pages 73–80, 2000.

[27] J. Dunn and C. Mayer. VARIATIONS: a digital music library system at indiana university. In *Proceedings of ACM Conference on Digital Libraries*, pages 12–19, 1999.

[28] S. Essid, G. Richard, and B. David. Musical instrument recognition based on class pairwise feature selection. In *Proceedings of the International Conference on Music Information Retrieval*, pages 560–567, 2003.

[29] E. Ferrari and G. Haus. The musical archive information system at Teatro alla Scala. In *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, volume II, pages 817–821, 1999.

[30] J.T. Foote. A similarity measure for automatic audio classification. In *Proceedings of AAAI Spring Symposium on Intelligent Integration and Use of Text, Image, Video, and Audio Corpora*, 1997.

[31] T. Fujishima. Realtime chord recognition of musical sound: a system using common lisp Music. In *Proceedings of the International Computer Music Conference*, pages 464–467, 1999.

[32] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, and D. Dobkin. A search engine for 3D models. *ACM Transactions on Graphics*, 22(1):83–105, 2003.

[33] A. Ghias, J. Logan, D. Chamberlin, and B.C. Smith. Query by humming: musical information retrieval in an audio database. In *Proceedings of the ACM Conference on Digital Libraries*, pages 231–236, 1995.

[34] P. Giannopoulos and R.C. Veltkamp. A pseudo-metric for weighted point sets. In *Proceedings of the European Conference on Computer Vision*, pages 715–730, 2002.

[35] E. Gómez and P. Herrera. Estimating the tonality of polyphonic audio files: cognitive versus machine learning modelling strategies. In *Proceedings of the International Conference on Music Information Retrieval*, pages 92–95, 2004.

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

[37] M. Goto and Y. Muraoka. An audio-based real-time beat tracking system and its applications. In *Proceedings of the International Computer Music Conference*, pages 17–20, 1998.

[38] F. Gouyon and S. Dixon. Dance music classification: a tempo-based approach. In *Proceedings of the International Conference on Music Information Retrieval*, pages 501–504, 2004.

Full text available at: http://dx.doi.org/10.1561/1500000002

[39] J.M. Grey and J.A. Moorer. Perceptual evaluations of synthesized musical instruments tones. *Journal of Acoustic Society of America*, 62(2):454–462, 1977.

[40] M. Grimaldi, P. Cunningham, and A. Kokaram. A wavelet packet representation of audio signals for music genre classification using different ensemble and feature selection techniques. In *Proceedings of the ACM International Workshop on Multimedia Information Retrieval*, pages 102–108, 2003.

[41] L. Grubb and R.B. Dannenberg. A stochastic method of tracking a vocal performer. In *Proceedings of the International Computer Music Conference*, pages 301–308, 1997.

[42] J. Haitsma, M. van der Veen, T. Kalker, and F. Bruekers. Audio watermarking for monitoring and copy protection. In *Proceedings of the ACM Workshops on Multimedia*, pages 119–122, 2000.

[43] Harmonica. Accompanying action on music information in libraries, July 2006. http://projects.fnb.nl/harmonica/.

[44] C. Harte, M. Sandler, S. Abdallah, and E. Gómez. Symbolic representation of musical chords: a proposed syntax for text annotations. In *Proceedings of the International Conference on Music Information Retrieval*, pages 66–71, 2005.

[45] J. Harvell and C. Clark. Analysis of the quantitative data of system performance. Deliverable 7c, LIB-JUKEBOX/4-1049: Music Across Borders, July 2006. http://www.statsbiblioteket.dk/Jukebox/edit-report-1.html.

[46] G. Haus and E. Pollastri. A multimodal framework for music inputs. In *Proceedings of the ACM Multimedia Conference*, pages 282–284, 2000.

[47] Y. Hijikata, K. Iwahama, K. Takegawa, and S. Nishida. Content-based music filtering system with editable user profile. In *Proceedings of the ACM Symposium on Applied Computing*, pages 1050–1057, 2006.

[48] K. Hoashi, K. Matsumoto, and N. Inoue. Personalization of user profiles for content-based music retrieval based on relevance Feedback. In *Proceedings of the ACM International Conference on Multimedia*, pages 110–119, 2003.

[49] H.H. Hoos, K. Renz, and M. Görg. GUIDO/MIR—an experimental musical information retrieval system based on GUIDO music notation. In *Proceedings of the International Symposium on Music Information Retrieval*, pages 41–50, 2001.

[50] J.-L. Hsu, C.C. Liu, and A.L.P. Chen. Efficient repeating pattern finding in music databases. In *Proceeding of the International Conference on Information and Knowledge Management*, pages 281–288, 1998.

[51] Humdrum. The Humdram Toolkit: software for music research, July 2006. http://www.music-cog.ohio-state.edu/Humdrum/.

[52] D. Huron. *The Humdrum Toolkit: Reference Manual*. Center for Computer Assisted Research in the Humanities, Menlo Park, CA, 1995.

[53] N. Hu, R.B. Dannenberg, and A.L. Lewis. A probabilistic model of melodic similarity. In *Proceedings of the International Computer Music Conference*, pages 509–515, 2002.

[54] N. Hu, R.B. Dannenberg, and G. Tzanetakis. Polyphonic audio matching and alignment for music retrieval. In *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 185–188, 2003.

Full text available at: http://dx.doi.org/10.1561/1500000002

[55] N. Hu and R.B. Dannenberg. A comparison of melodic database retrieval techniques using sung queries. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, pages 301–307, 2002.

[56] IMIRSEL. The international music information retrieval system evaluation laboratory project, July 2006. http://www.music-ir.org/evaluation/.

[57] International Organization for Standardization (ISO). Information technology: Standard Music Description Language (SMDL). Technical Report ISO/IECDIS 10743:1995, ISO International Electrotechnical Commission, Geneva, Switzerland, July 2006. ftp://ftp.techno.com/pub/SMDL/.

[58] E. Isaacson. What you see is what you get: on visualizing music. In *Proceedings of the International Conference on Music Information Retrieval*, pages 389–395, 2005.

[59] ISMIR. The international conferences on music information retrieval, July 2006. http://www.ismir.net/.

[60] K. Sparck Jones and P. Willett. *Readings in Information Retrieval*. Morgan Kaufmann, San Francisco, CA, 1997.

[61] T. Kageyama and Y. Takashima. A melody retrieval method with hummed melody (written in japanese). *Transactions of the Institute of Electronics, Information and Communication Engineers*, J77D-II(8):1543–1551, 1994. Cited in Ghias *et al.*, 1995.

[62] A. Kapur, M. Benning, and G. Tzanetakis. Query-by-beat-boxing: music retrieval for the DJ. In *Proceedings of the International Conference on Music Information Retrieval*, pages 170–177, 2004.

[63] M.L. Kherfi, D. Ziou, and A. Bernardi. Image retrieval from the World Wide Web: issues, techniques, and systems. *ACM Computing Surveys*, 36(1):35–67, 2004.

[64] T. Kitahara, M. Goto, K. Komatani, T. Ogata, and H.G. Okuno. Instrument identification in polyphonic music: feature weighting with mixed sound, pitch dependent timbre modeling, and use of musical content. In *Proceedings of the International Conference on Music Information Retrieval*, pages 558–563, 2005.

[65] R. Kochumman, C. Monroy, R. Furuta, A. Goenka, E. Urbina, and E. Melgoza. Towards an electronic variorum edition of Cervantes' Don Quixote: visualizations that support preparation. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, pages 199–200, 2002.

[66] C.L. Krumhansl. Why is musical timbre so hard to understand? In S. Nielsen and O. Olsson, editors, *Structure and Perception Electroacoustic Sound and Music*, pages 45–53. Elsevier, Amsterdam, NL, 1989.

[67] C.L. Krumhansl. The geometry of musical structure: a brief introduction and history. *Computer in Entertainment*, 3(4):1–14, 2005.

[68] F.-F. Kuo and M.-K. Shan. Looking for new, not known music only: music retrieval by melody style. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, pages 243–251, 2004.

[69] V. Lavrenko and J. Pickens. Polyphonic music modeling with random fields. In *Proceedings of the ACM International Conference on Multimedia*, pages 120–129, 2003.

Full text available at: http://dx.doi.org/10.1561/1500000002

[70] G.H. Leazer. The effectiveness of keyword searching in the retrieval of musical works on sound recordings. *Cataloging and Classification Quarterly*, 15(3):15–55, 1992.

[71] J.H. Lee and J.S. Downie. Survey of music information needs, uses, and seeking behaviours: preliminary findings. In *Proceedings of the International Conference on Music Information Retrieval*, pages 441–446, 2004.

[72] F. Lerdhal and R. Jackendoff. *A Generative Theory of Tonal Music*. The MIT Press, Cambridge, MA, 1983.

[73] M. Lesaffre, M. Leman, K. Tanghe, B. De Baets, H. De Meyer, and J.-P. Martens. User-dependent taxonomy of musical features as a conceptual framework for musical audio-mining technology. In *Proceedings of the Stockholm Music Acoustics Conference*, pages 635–638, 2003.

[74] M. Lesaffre, K. Tanghe, G. Martens, D. Moelants, M. Leman, B. De Baets, H. De Meyer, and J.-P. Martens. The MAMI query-by-voice experiment: collecting and annotating vocal queries for music information retrieval. In *Proceedings of the International Conference on Music Information Retrieval*, pages 65–71, 2003.

[75] G. Linden, B. Smith, and J. York. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1):76–80, 2003.

[76] N.-H. Liu, Y.-H. Wu, and A.L.P. Chen. Efficient K-NN search in polyphonic music databases using a lower bounding mechanism. In *Proceedings of the ACM International Workshop on Multimedia Information Retrieval*, pages 163–170, 2003.

[77] Q. Li, B.M. Kim, D.H. Guan, and D.W. Oh. A music recommender based on audio features. In *Proceedings of the ACM International Conference on Research and Development in Information Retrieval (SIGIR)*, pages 532–533, 2004.

[78] T. Li, M. Ogihara, and Q. Li. A comparative study on content-based music genre classification. In *Proceedings of the ACM International Conference on Research and Development in Information Retrieval (SIGIR)*, pages 282–289, 2003.

[79] B. Logan. Content-based playlist generation: exploratory experiments. In *Proceedings of the International Conference on Music Information Retrieval*, pages 295–296, 2002.

[80] B. Logan. Music recommendation from song sets. In *Proceedings of the International Conference on Music Information Retrieval*, pages 425–528, 2004.

[81] D. Lübber. Sonixplorer: combining visualization and auralization for content-based exploration of music collections. In *Proceedings of the International Conference on Music Information Retrieval*, pages 590–593, 2005.

[82] A. Lubiw and L. Tanur. Pattern matching in polyphonic music as a weighted geometric translation problem. In *Proceedings of the International Conference of Music Information Retrieval*, pages 289–296, 2004.

[83] MAMI. Musical Audio Mining—"query by humming", July 2006. http://www.ipem.ugent.be/MAMI/.

[84] S. Mcadams. Perspectives on the contribution of timbre to musical structure. *Computer Music Journal*, 23(3):85–102, 1999.

Full text available at: http://dx.doi.org/10.1561/1500000002

[85] A. McLane. Music as information. In M.E. Williams, editor, *Arist*, volume 31, chapter 6, pages 225–262. American Society for Information Science, 1996.
[86] C. Meek and W. Birmingham. Johnny can't sing: a comprehensive error model for sung music queries. In *Proceedings of the International Conference on Music Information Retrieval*, pages 65–71, 2002.
[87] C. Meek and W. Birmingham. Automatic thematic extractor. *Journal of Intelligent Information Systems*, 21(1):9–33, 2003.
[88] M. Melucci, N. Orio, and M. Gambalunga. An evaluation study on music perception for musical content-based information Retrieval. In *Proceedings of the International Computer Music Conference*, pages 162–165, 2000.
[89] M. Melucci and N. Orio. Musical information retrieval using melodic surface. In *Proceedings of the ACM Conference on Digital Libraries*, pages 152–160, 1999.
[90] M. Melucci and N. Orio. Combining melody processing and information retrieval techniques: methodology, evaluation, and system implementation. *Journal of the American Society for Information Science and Technology*, 55(12):1058–1066, 2004.
[91] R. Middleton. *Studying Popular Music*. Open University Press, Philadelphia, PA, 2002.
[92] Mirex 2006 Wiki. Second annual music information retrieval evaluation exchange, July 2006. http://www.music-ir.org/mirex2006/.
[93] F. Mörchen, A. Ultsch, M. Nöcker, and C. Stamm. Databionic visualization of music collections according to perceptual distance. In *Proceedings of the International Conference on Music Information Retrieval*, pages 396–403, 2005.
[94] MPEG. The MPEG home page, July 2006. http://www.chiariglione.org/mpeg/.
[95] M. Müller, F. Kurth, and M. Clausen. Audio matching via chroma-based statistical features. In *Proceedings of the International Conference of Music Information Retrieval*, pages 288–295, 2005.
[96] MuseData. An electronic library of classical music scores, July 2006. http://www.musedata.org/.
[97] Musica. The international database of choral repertoire, July 2006. http://www.musicanet.org/.
[98] MusicXML. Recordare: Internet music publishing and software, July 2006. http://www.musicxml.org/.
[99] GUIDO Music Notation. The GUIDO NoteServer, July 2006. http://www.noteserver.org/.
[100] Music Notation. Formats, July 2006. http://www.music-notation.info/.
[101] MusiXTeX. MusiXtex and related software, July 2006. http://icking-music-archive.org/software/indexmt6.html.
[102] E. Narmour. *The Analysis and Cognition of Basic Melodic Structures*. University of Chicago Press, Chicago, MI, 1990.
[103] G. Neve and N. Orio. Indexing and retrieval of music documents through pattern analysis and data Fusion Techniques. In *Proceedings of the International Conference on Music Information Retrieval*, pages 216–223, 2004.

Full text available at: http://dx.doi.org/10.1561/1500000002

[104] N.M. Norowi, S. Doraisamy, and R. Wirza. Factors affecting automatic genre classification: an investigation incorporating non-western musical forms. In *Proceedings of the International Conference on Music Information Retrieval*, pages 13–20, 2005.

[105] N. Orio and G. Neve. Experiments on segmentation techniques for music documents indexing. In *Proceedings of the International Conference on Music Information Retrieval*, pages 104–107, 2005.

[106] N. Orio. Alignment of performances with scores aimed at content-based music access and Retrieval. In *Proceedings of European Conference on Digital Libraries*, pages 479–492, 2002.

[107] R.P. Paiva, T. Mendes, and A. Cardoso. On the detection of melody notes in polyphonic audio. In *Proceedings of the International Conference on Music Information Retrieval*, pages 175–182, 2005.

[108] E. Pampalk, S. Dixon, and G. Widmer. Exploring music collections by browsing different views. In *Proceedings of the International Conference on Music Information Retrieval*, pages 201–208, 2003.

[109] C.L. Parker. A tree-based method for fast melodic retrieval. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, pages 254–255, 2004.

[110] G. Peeters. Rhythm classification using spectral rhythm patterns. In *Proceedings of the International Conference on Music Information Retrieval*, pages 644–647, 2005.

[111] J. Pickens and T. Crawford. Harmonic models for polyphonic music retrieval. In *Proceedings of the International Conference on Information and Knowledge Management*, pages 430–437, 2002.

[112] A. Pienimäki and K. Lemström. Clustering symbolic music using paradigmatic and surface level analyses. In *Proceedings of the International Conference of Music Information Retrieval*, pages 262–265, 2004.

[113] A. Pienimäki. Indexing music database using automatic extraction of frequent phrases. In *Proceedings of the International Conference on Music Information Retrieval*, pages 25–30, 2002.

[114] S. Della Pietra, V. Della Pietra, and J. Lafferty. Inducing features of random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:380–393, 1997.

[115] A. Pikrakis, I. Antonopoulos, and S. Theodoris. Music meter and tempo tracking from raw polyphonic audio. In *Proceedings of the International Conference on Music Information Retrieval*, pages 192–197, 2004.

[116] L. Rabiner and B.-H. Juang. *Fundamentals of Speech Recognition*. Prentice-Hall, Englewood Cliffs, NJ, 1993.

[117] C. Raphael and J. Stoddard. Harmonic analysis with probabilistic graphical models. In *Proceedings of the International Conference on Music Information Retrieval*, pages 177–181, 2003.

[118] C. Raphael. Automatic segmentation of acoustic musical signals using hidden markov models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4):360–370, 1999.

[119] T. Rohdenburg, V. Hohmann, and B. Kollmeier. Objective perceptual quality measures for the evaluation of noise reduction schemes. In *Proceedings of the*

Full text available at: http://dx.doi.org/10.1561/1500000002

International Workshop on Acoustic Echo and Noise Control, pages 169–172, 2005.

[120] J. Rothstein. *Midi: A Comprehensive Introduction*. A-R Editions, Madison, WI, 1991.

[121] E. Selfridge-Field. *Beyond MIDI: The Handbook of Musical Codes*. MIT Press, Cambridge, MA, 1997.

[122] Shazam. Home page, July 2006. http://www.shazam.com/.

[123] J. Shifrin, B. Pardo, C. Meek, and W. Birmingham. Hmm-based musical query retrieval. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, pages 295–300, 2002.

[124] H.-H. Shih, S.S. Narayanan, and C.-C.J. Kuo. Multidimensional humming transcription using hidden Markov models for query by humming systems. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 541–544, 2003.

[125] D.R. Stammen and B. Pennycook. Real-time recognition of melodic fragments using the dynamic timewarp algorithm. In *Proceedings of the International Computer Music Conference*, pages 232–235, 1993.

[126] R. Stenzel and T. Kamps. Improving content-based similarity measures by training a collaborative model. In *Proceedings of the International Conference on Music Information Retrieval*, pages 264–271, 2005.

[127] MIR Systems. A survey of music information retrieval systems, July 2006. http://mirsystems.info/.

[128] J. Tenney and L. Polansky. Temporal gestalt perception in music. *Journal of Music Theory*, 24(2):205–241, 1980.

[129] P. Toiviainen. Visualization of tonal content with self-organizing maps and self-similarity Matrices. *Computers in Entertainment*, 3(4):1–10, 2005.

[130] TREC. Text REtrieval Conference home page, July 2006. http://trec.nist.gov/.

[131] Y. Tseng. Content-based retrieval for music collections. In *Proceedings of the ACM International Conference on Research and Development in Information Retrieval (SIGIR)*, pages 176–182, 1999.

[132] R. Typke, M. den Hoed, J. de Nooijer, F. Wiering, and R.C. Veltkamp. A ground truth for half a million musical incipits. *Journal of Digital Information Management*, 3(1):34–39, 2005.

[133] R. Typke, R.C. Veltkamp, and F. Wiering. Searching notated polyphonic music using transportation distances. In *Proceedings of the ACM International Conference on Multimedia*, pages 128–135, 2004.

[134] R. Typke, R.C. Veltkamp, and F. Wiering. Evaluating retrieval techniques based on partially ordered ground truth lists. In *Proceedings of the International Conference of Multimedia and Expo*, 2006.

[135] R. Typke, F. Wiering, and R.C. Veltkamp. A survey on music information retrieval systems. In *Proceedings of the International Conference on Music Information Retrieval*, pages 153–160, 2005.

[136] G. Tzanetakis and P. Cook. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5):293–302, 2002.

Full text available at: http://dx.doi.org/10.1561/1500000002

[137] A. Uitdenbogerd and J. Zobel. Manipulation of music for melody matching. In *Proceedings of the ACM Conference on Multimedia*, pages 235–240, 1998.

[138] E. Ukkonen, K. Lemström, and V. Mäkinen. Geometric algorithms for transposition invariant content-based music retrieval. In *Proceedings of the International Conference of Music Information Retrieval*, pages 193–199, 2003.

[139] E. Unal, S.S. Narayanan, and E. Chew. A statistical approach to retrieval under user-dependent uncertainty in query-by-humming systems. In *Proceedings of the ACM International Workshop on Multimedia Information Retrieval*, pages 113–118, 2004.

[140] E. Vincent and X. Rodet. Instrument identification in solo and ensemble music using independent subspace analysis. In *Proceedings of the International Conference on Music Information Retrieval*, pages 576–581, 2003.

[141] Wedelmusic. Server of sound and sound-processing, July 2006. http://www.wedelmusic.org.

[142] G.A. Wiggins, K. Lemström, and D. Meredith. SIA(M)ESE: an algorithm for transposition invariant, polyphonic content-based music retrieval. In *Proceedings of the International Conference of Music Information Retrieval*, pages 283–284, 2002.

[143] E. Wold, T. Blum, D. Keislar, and J. Wheaton. Content-based classification, search, and retrieval of audio. *IEEE Multimedia*, 3(3):27–36, 1996.

[144] T. Zhang and C.-C. Jay Kuo. Hierarchical system for content-based audio classification and retrieval. In *Proceedings of International Conference on Speech, Audio, and Signal Processing, Vol. 6*, pages 3001–3004, 1999.