The Exploratory Workshop on Music Information Retrieval

19 August 1999
ACM SIGIR '99
Berkeley, California

Proceedings Compiled by
J. Stephen Downie
Information Retrieval Group
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
jdownie@uiuc.edu

Introduction

This workshop is intended to bring together Music Information Retrieval (MIR) researchers; Information Retrieval (IR) researchers; computer scientists; musicologists; music bibiliographers; music, digital, and general librarians; and, music providers and music industry members; in a forum exclusively devoted to issues pertaining to MIR. We are conducting this workshop in order achieve the following objectives.

Objectives

1. Foster the development of methods and technologies for content-based Music Information Retrieval (MIR) systems by:

a. having those on the user side (i.e, musicologists, librarians, industry members, etc.) inform those on the research side (i.e., IR researchers, computer scientists, etc.) about real-world problems, needs, and opportunities

b. having those on the research side inform those on the user side about the strengths and weaknesses of the various approaches to content-based MIR (both extant and under research & development)

c. having those with a specific research interest in MIR demonstrate to other IR researchers the interesting, unique, and non-trivial research issues involved in the development of MIR systems

2. Foster a framework for future fruitful research into MIR by:

a. having participants explore consensus opinion on the establishment of research priorities, inter-disciplinary collaborations, evaluation standards, test collections, resource sharing, funding opportunities, communication channels, etc.

b. having participants begin work on the establishment of a formal organization of those interested in MIR. Such an organization will have as its mandate those items mentioned above.

Workshop Timetable

Exploratory Workshop on Music Information Retrieval 19 August 1999

Prepared by J. Stephen Downie
jdownie@uiuc.edu

N.B The Q & A times are at the discretion of the presenters. There is time set aside (Response) for questions should the talks take up the full twenty minutes.

Code	Name
CMN	C. Nevill -Manning
DG	D. Gibson
DH	D. Huron
ESF	E. Selfridge-Field
GB	G.Bowles
GF, GH & EP	G. Frazinni, G. Haus & E. Pollastri
JSD	J. Stephen Downie
MD	M. Dovey
MF	M. Fingerhut
MKD	M. K. Duggan
MM & NO	M. Melucci & N. Orio
TH	T. Heck

Morning Session			Moderator
9:00	9:04	*INTRO*	
9:05	9:09	*INTRO*	
9:10	9:14	**MKD**	D
9:15	9:19		O
9:20	9:24		W
9:25	9:29	*Q & A*	N
9:30	9:34	**CNM**	I
9:35	9:39		E
9:40	9:44		
9:45	9:49	*Q & A*	
9:50	9:54	**GF, GH & EP**	
9:55	9:59		
10:00	10:04		
10:05	10:09	*Q & A*	
10:10	10:14	**MD**	
10:15	10:19		
10:20	10:24		
10:25	10:29	*Q & A*	
10:30	10:34	*Response*	
10:35	10:39	*Response*	
10:40	10:44	*Coffee Break*	
10:45	10:49	*Coffee Break*	
10:50	10:54	**DG**	H
10:55	10:59		U
11:00	11:04		R
11:05	11:09	*Q & A*	O
11:10	11:14	**JSD**	N
11:15	11:19		
11:20	11:24		
11:25	11:29	*Q & A*	
11:30	11:34	**MM & NO**	
11:35	11:39		
11:40	11:44		
11:45	11:49	*Q & A*	
11:50	11:54	*Response*	
11:55	11:59	*Response*	
12:00			

Afternoon Session			Moderator
13:00	13:04	*INTRO*	
13:05	13:09	**ESF[Keynote]**	N
13:10	13:14		E
13:15	13:19		V
13:20	13:24	*Q & A*	I
13:25	13:29	**MF**	L
13:30	13:34		L
13:35	13:39		
13:40	13:44	*Q & A*	M
13:45	13:49	**TH**	A
13:50	13:54		N
13:55	13:59		N
14:00	14:04	*Q & A*	I
14:05	14:09	**GB**	N
14:10	14:14		G
14:15	14:19		
14:20	14:24	*Q & A*	
14:25	14:29	*Response*	
14:30	14:34	*Response*	
14:35	14:39	*Coffee Break*	
14:40	14:44	*Coffee Break*	
14:45	14:49	**DH**	D
14:50	14:54		O
14:55	14:59		W
15:00	15:04	*Q & A*	N
15:05	15:09	Organization Discussion	I
15:10	15:14		E
15:15	15:19		
15:20	15:24		
15:25	15:29		
15:30	15:34		
15:35	15:39		
15:40	15:44		
15:45	15:49		
15:50	15:54		
15:55	15:59		
16:00		Possibly go later.	

Table of Abstracts

Test Bed: A Virtual Library of 19th-Century Sheet Music

Abstract

The California Sheet Music Project (www.sims.berkeley.edu/~mkduggan/neh.html) includes images of some 2,000 editions 10,000 images of 19th-century California sheet music, rare material on fragile 19th-century paper, with associated sound and video files. The Project was designed to be an image test bed for extending the capacity of MARC (Machine-Readable Cataloging, the standard digital format for library catalogues) for multimedia description and retrieval through the MARC/SGML/HTML search engine of Ray Larson (Cheshire II). The funding source, the California State Library (federal funds [1997-1998; 1999] through the Library Services and Technology Act, Title IIB, demanded a retrieval product capable of searching descriptive alphabetic information through established international bibliographic networks in the MARC format to retrieve music images. The World Wide Web and the MARC field for web links (856) have become robust enough to carry links for sound and video, allowing interaction of modalities and multimodal query formation. The challenge for information retrieval of the test bed of digitized sheet music lies in the importance of the material as **(1) music (songs, piano, guitar and other instrumental music), (2) art (covers with engravings, woodcuts, lithographs, attached photographs), (3) multicultural texts, (4) sound and video, (5) regional history,** and **(6) printing and publishing history** (engraved plates, hot type, plate processes), each important to a constituency.

1. Music Part of the grant proposal asked for funding to encode incipits of song verses and refrains for digital retrieval using the system of grant associate Garrett Bowles (Music, UCSD). Currently only unsearchable scanned images of the music are available, with access to alphabetic information (composer, title, date, instrument, etc.).
2. Art By August, 1999, retrieval of the full depth of 400 dpi TIFF images becomes available through the JTIP (Jpeg Tiled Image Pyramid) image file format Project Associate MIP [Museum Informatics Project (www.mip.berkeley.edu)] and Netimage (www.avelem.fr). The original large format of published music maintained in the archival digital files is available through JTIP for zoom review on a monitor, and full-size printing for performance. MARC coding using the Library of Congress Thesaurus of Graphics Materials for subjects and genre categories for types of artistic processes allow retrieval of material pictured on the sheet music covers. MARC authority records allow retrieval of all variant forms of associated names (portraits of composers and performers, artists, printers) and organizations (buildings, city scenes, parks, ranches, advertising products).
3. Multicultural Texts A major reason for the funding of the project was the importance of the music for the multicultural history of the West, including alphabetic texts in many languages.
4. Sound and Video As sound recordings and videos associated with the music of the project come to light, they are digitized in segments matching the printed music and entered as web links in the MARC records. The multimodal cataloging rules of MARC allow multimodal query formation. Thus one record might be retrievable in image, MIDI sound, analog sound, and video, with subject and genre access specific to a medium. MARC has only recently been extended to allow inclusion of URLs of "related resources" and "versions of resources."
5-6. Regional History. Printing and Publishing History. Authorized forms of names, places, events, subjects.

Author Information

Mary Kay Duggan
Music; School of Information Management and Systems
University of California, Berkeley
102 South Hall, Berkeley, CA 94720-4600
510-642-5764; Fax: 510-643-5323; mkduggan@sims.berkeley.edu

Suggested Readings

Bowles, Garrett. *Thematic Catalog of the Works of Marin Marais.* Http:orpheus.ucsd.edu/music/ghb.html

Duggan, Mary Kay. ACalifornia Sheet Music Project.@ *RLIN Focus* Issue 29 (December 1997), 8-9.

A856 Electronic Location and Access.@ In *US MARC Concise Format for Bibliographic Data.* Library of Congress, Network Development and MARC Standards Office, 1999. Http://lcweb.loc.gov/marc/bibliographic/ecbdhome.htm

A Digital Library of Popular Music

Craig G. Nevill-Manning

My talk describes a project in its infancy—a large-scale digital library of popular music. We currently have pilot system providing a query-by-singing interface to 1,000 tunes, and a larger system providing full-text retrieval of 100,000 tunes based on metadata. Our goal is a digital library that people in all walks of life will want to use. Popular music culture tends to transcend social boundaries, be they national borders or socioeconomic groupings. Yet music simultaneously exhibits strong cultural diversity: folk music tends to be specific to a particular country or region, and different styles even define ethnic groupings within a country. Because people can often remember a tune but not necessarily its title or composer, our digital library will be able to retrieve a tune based on the user humming a short excerpt.

A crucial issue is source material. Our digital library is based on Musical Instrument Digital Interface (MIDI) format files, which provide a symbolic representation of music. MIDI files allow computation to be performed directly on the sequence of notes in a tune, in contrast with sampled music, such as CD tracks compressed using the MP3 technique, where individual notes are not made explicit. MIDI files are widely available on the Web. Indeed, the mixture of musical genres that are available on the Web reflects the mixture of musical styles in rough proportion to its general popularity amongst educated Web users. MIDI files can easily be transmitted at low bandwidth, encouraging portability and hence usability.

We currently have an unsophisticated approach to extracting the melody from a MIDI file and comparing a melody to a sung query. We plan to refine these techniques and scale them up. Search is currently performed by exhaustive pairwise comparison between the query and all the tunes in the database. To accelerate this process, we will draw inspiration from bioinformatics, where techniques for fast retrieval of protein sequences have been in use for several years. Music similarity computations resemble protein similarity computations because both must be robust to deletions, insertions and substitutions. The BLAST technique for protein search operates by creating an inverted index of n-grams and using it to exclude tunes that are unlikely to be highly similar to the query. Full dynamic programming is performed on the remaining candidates, which are usually in the order to a few hundred. Finally, we plan to provide automatic clustering to recognize multiple versions of a song, and develop machine learning approaches to classify tunes according to genre.

Cross Automatic Indexing of Score and Audio Sources: Approaches for Music Archive Applications

Abstract

In this talk we present the latest research efforts made by L.I.M. ("Musical Informatics Laboratory") in the field of Music Information Retrieval, Automatic Music Analysis, Automatic Audio Understanding and the cross relations existing between them. These topics are deeply investigated in the framework of the "Music Archive Project" at Teatro alla Scala which is continuously under development since 1996.

Developing a multimedia database for this music archive poses several interesting challenges, due to the complexity of the database schema and to the unstructured properties of multimedia data (Haus 1998). For this reason, the acquisition methods and the content processing of audio and score data were deeply investigated. The OMR (Optical Music Recognition) performance is improved with a tool that deskews the staves, normalizes the number of staves per system and widens the closest staves. Moreover, translator tools from/to NIFF, SMDL, MIDI, ENIGMA formats are available. A more interesting task from the research point of view is the segmentation of a score; this means finding the "theme" in a music piece. It can be done by recursively calculating a set of melodic and rhythmic operators applied to the symbolic representation of the score (i.e. NIFF format file) and measuring the distance between a theme hypothesis and all its repetitions (filtered through the same set of operators) along the score. Next to the textual query, we developed two melody retrieval systems based on approximate string matching algorithm (Pollastri 1998) and on morphological musical metrics (Polanski 1996). Both can be applied to an acoustic input or a symbolic input, i.e. a score fragment. A comparative analysis of the effectiveness of these algorithms is currently under development.

The audio source needs to be organized as well; this problem can be addressed as "audio segmentation". At a first level, this means identifying the starting and ending points of a recording; in other words, every audio file must be segmented into silence (or noise and applause), speech and music. With the recently approved MPEG4 (Structured-Audio) standard, a software tool capable of segmenting automatically an audio source is by no means very important; nevertheless, we are able to do that only for very simple and monophonic sources. In the case of music, the existence of a symbolic description of the audio content – its score – could be very helpful. In fact, one of the key features of a complete digital music archive is the cross-relation between the recording of a music performance and its score. A musically educated user can move forward and backward on a recording keeping the right position in a score and vice versa. We want to achieve this automatically. Since a blind recognition of music-audio sources is still too difficult, we propose to use the score as "indexing template" for audio segmentation. This knowledge is used to lead us in an audio analysis score driven. The ultimate target is a list of markers that binds every acoustic event with its symbolic counterpart on the score. With this feature, a new degree of freedom is achieved by the system: the user could look for a score or a digitalized recording without any restriction. From the point of view of the database, every performance is represented by some textual information (author, title, date, musicians, director,…), the scores played (in any of the supported format) and a table of pointers to the CDs containing the digitalized recordings extended with the list of markers that binds the performance with the right points in the scores. The resulting database architecture is enough powerful and complete to show that an integrated approach for both audio and score sources is now concretely possible; this moves the music retrieval systems toward the most general framework of the multimedia databases.

Author Information

Giuseppe Frazzini, Goffredo Haus, Emanuele Pollastri
L.I.M./D.S.I.-Dept. of Computer Science
State University of Milan
via Comelico, 39 / I-20135 Milano (Italy)
Tel: +39 02 55006222 / Fax: +39 02 55006373
giusfra@globalnet.it, haus@dsi.unimi.it, pollastri@dsi.unimi.it

Suggested Readings

Haus, Goffredo. 1998. Rescuing La Scala's Audio Archives. *IEEE Computer* 31(3): 88-89.

Pollastri, Emanuele. 1998. Melody Retrieval Based on Approximate String Matching and Pitch-Tracking Methods. *XII Colloquium of Musical Informatics, AIMI/Univesity of Udine, Gorizia*, 1998.

Polanski, Larry. 1996. Morphological Metrics. *Journal of New Music Research* 25(4): 289-368.

OMRAS - NSF/JISC Funded International Collaborative Digital Library Project for Online Music Search and Retrieval

Abstract

Retrieving information from multimedia archives in general is hampered by the lack of available tools for searching their specialized content. This project aims to address one area of this problem, that of searching musical databases. The music in the databases might be stored in a variety of formats and comprise any of three basic types of data:

a) encoded descriptive score-data (as in the files of a commercial score-notation program);
b) listings of prescriptive performance instructions (as in MIDI files);
c) sound-files (as in digital audio recordings).

OMRAS is one of six international digital library projects jointly funded by the NSF in the US and JISC in the UK, and is a collaborative project of the Center for Intelligent Information Retrieval at the University of Massachusetts, and Kings College London. The 3-year project will focus principally on four aspects:

1. System architecture, internal data-format, data-conversion;
2. Preprocessing and indexing of musical data, implementation of efficient search algorithms;
3. User interfaces for search-queries and for result-presentation;
4. Audio music-recognition sufficient to generate data in the internal format.

One aspect of OMRAS is the searching and indexing of polyphonic music, and there will be a demonstration of some preliminary work in this area.

Author Information

Matthew J. Dovey	Tim Crawford	Don Byrd
Department of Computer Science	Department of Music	Center for Intelligent Information Retrieval
Kings College London	Kings College London	Computer Science Department
Strand	Strand	University of Massachusetts
London WC2R 2LS, U.K.	London WC2R 2LS, U.K.	Amherst, MA 01003
Matthew.Dovey@las.ox.ac.uk	Tim.Crawford@kcl.ac.uk	dbyrd@cs.umass.edu

Suggested Readings

D. Bainbridge. 1998. MELDEX: A Web-based Melodic Locator Service. *Computing in Musicology.* 11: 223-229.

I. V. Bakhmutova, V. D. Gusev and T. N. Titkova. 1997. The Search for Adaptations in Song Melodies. *Computer Music Journal*. 12(1): 58-67.

E. Cambouropoulos. 1996. A General Pitch Interval Representation: Theory and Applications. *Journal of New Music Research.* 25 (3): 231-251.

T. C. C. and A. L. P. Chen. 1998. Query by Rhythm: An Approach for Song Retrieval in Music Databases. *Proceedings of IEEE International Workshop on Research Issues in Data Engineering.* 139-146.

T. Crawford, C. S. Iliopoulos, R. and Raman. 1998. String Matching Techniques for Musical Similarity and Melodic Recognition. *Computing in Musicology.* 11: 73-100.

M. Dovey. 1999. A matrix based algorithm for locating polyphonic phrases within a polyphonic musical piece. *Focus Workshop on Pattern Processing in Music Analysis and Creation. Symposium on Artificial Intelligence and Musical Creativity. Convention for the Society for the Study of Artificial Intelligence and Simulation of Behaviour.*

A. Körnstadt. 1998. ThemeFinder: A Web-based Melodic Search Tool. *Computing in Musicology*. 11: 231-236.

P. McGettrick. 1997. MIDIMatch: Musical Pattern Matching in Real Time. MSc Dissertation, York University, U.K.

E. Selfridge-Field, (ed). 1997. *Beyond MIDI: The handbook of musical codes*. MIT Press, Cambridge, Mass.

E. Selfridge-Field. 1998. Conceptual and Representational Issue in Melodic Comparison. *Computing in Musicology*. 11: 3-64.

L. A. Smith, R. J. McNab and I. H. Witten. 1998. Sequence-based Melody Comparison: A Dynamic-Programming Approach. *Computing in Musicology*. 11: 101-118.

Name That Clip

Abstract

This is a system which is capable of identifying a song, given any short clip from it, such as a five to ten second radio sample. In contrast to "name that tune" approaches, it does not attempt to identify the main melody, but rather tries to extract a summary of the sound stream on a more basic level. It assumes that the sample is no more than a rerecording of the original: specifically, the system is designed to be immune to white noise and band-pass filtering. As such, it has uses as a public service for consumers and music suppliers, or as a radio monitoring service for royalty protection.

The basic technique is to subdivide the audio stream into discrete intervals, and extract a packet of summary bits from each interval. The same summarization is applied to the query sample, and lookup is done by a form of approximate substring matching algorithm. Since the subdivisions of the query sample will not be in phase with the subdivisions in the database, the query is resubdivided at several different offsets, and the best match is found over all offsets. Each interval is summarized by examining the changes in the frequency spectrum from that interval to the following interval: The spectrum is partitioned into bins, and a particular summary bit is set when the frequency transitions from one particular bin to another.

We can thus build up a large database of songs, as this summarization yields about 100 bits per second of music. The indexing task is to search for the database sequence with the closest Hamming distance to the query. This is done by a locality-preserving hashing technique.

The system has been tested to date on a database of 1500 songs, spanning a variety of genres, and with a query set of 500 microphone-recorded samples, with a near-perfect success rate.

Author Information

David Gibson
Department Of Computer Science
UC Berkeley
581 Soda Hall
Berkeley, CA, 94720-1776
Home: (510) 841-9068 Office: (510) 643-5425
dag@cs.berkeley.edu

Suggesting Readings

Further information and demos will be available at:
http://www.cs.berkeley.edu/~dag/NameThatClip/

Applying Text Retrieval Methods to Music Information Retrieval

Abstract

Taking our cue from those printed thematic catalogues that have reduced the amount of music information represented we developed, and then evaluated, a MIR system based upon the intervals found within the melodies of a collection of 9354 folksongs. We believe that there is enough information contained within an interval-only representation of monophonic melodies that effective retrieval of music information has been achieved. We extended the thematic catalogue model by affording access to musical expressions found anywhere within a melody. To achieve this extension we fragmented to the melodies into length-n subsections called n-grams. The length of these n-grams and the degree to which we precisely represent the intervals are variables analyzed in this study.

N-grams form discrete units of melodic information much in the same manner as words are discrete units of language. Thus, we have come to consider them "musical words." This implies that, for the purposes of music information retrieval, we can treat them as "real words" and thereby apply traditional text-based information retrieval techniques. We examined the validity of our "musical word" concept in two ways. First, a variety of informetric analyses were conducted to examine in which ways the informetric properties of "musical words" and "real words" are similar or different. Second, we constructed a collection of "musical word" databases using the famous text-based SMART information retrieval system. A group of simulated queries were run against these databases. The results were evaluated using the normalized precision and normalized recall measures. Results indicate that the simple approach to music information retrieval examined in this study shows great merit.

Author Information

J. Stephen Downie
University of Illinois at Urbana-Champaign
Information Retrieval Group
Graduate School of Library and Information Science
501 East Daniel St.
Champaign, IL 61820
jdownie@uiuc.edu

Suggested Readings

Downie, J. Stephen. 1999a. *Evaluating a Simple Approach to Music Information Retrieval: Conceiving Melodic N-grams as Text*. Ph.D. Thesis. London, Ontario: University of Western Ontario.

Downie, J. Stephen. 1999b . Music retrieval as text retrieval: Simple yet effective. In *Proceedings of the Association for Computing Machinery, SIGIR '99 conference, University of California at Berkeley, 15-19 August 1999, Berkeley, California*. New York: Association for Computing Machinery.

Downie, J. Stephen. 1998. Informetrics and music information retrieval: An informetric examination of a folksong database. In *Information Science at the Dawn of the New Millennium. Proceedings of the 26th Annual Conference of the Canadian Association for Information Science, 3-5 June 1998, Ottawa, Ontario,* 375-392. Toronto: Canadian Association for Information Science.

Downie, J. Stephen. 1997. Informetrics and music information retrieval. In Communication and Information in Context: Society, Technology, and the Professions. Proceedings of the 25th Annual Conference of the Canadian Association for Information Science, 8-10 June 1997, St. John's, Newfoundland, 295-308. Toronto: Canadian Association for Information Science.

Downie, J. Stephen. 1995. The MusiFind Music Information Retrieval Project, Phase III: Evaluation of indexing options. In *Connectedness: Information, systems, people, organizations, Proceedings of the 23rd Annual Conference of the Canadian Association for Information Science, 7-10 June 1995, Edmonton, Alberta,* 135-146. Toronto: Canadian Association for Information Science.

The use of melodic surface segmentation to retrieve musical information through querying and browsing automatically constructed hypermusic

Abstract

The basic idea of this work is that the standard principles of textual information retrieval (IR) can be applied to music retrieval. The methodology is based on the following steps: 1) extraction of melodies from polyphonic musical scores or queries in MIDI format; 2) segmentation of melodies in musical phrases, where the concept of "phrase" corresponds to the concept of "term" in textual IR; 3) interval quantization and duration normalization of musical phrases; 4) index phrase generation and weighting; 5) application of the vector-space model to retrieve musical documents against musical queries; 6) automatic construction of hypermusic, where by hypermusic we mean a set of musical documents or phrases connected through a network of hyperlinks which can be both queried and navigated.

The algorithm for the segmentation of the melody in "lexical units", that is in musical phrases, is based on the Local Boundaries Detection Model (Cambouropoulus 1997). In this model it is proposed that each melody has some points where listeners perceive the presence of a change in the melodic flown. A given weight is assigned to each note, depending on the melodic structure, and in particular depending on the relationship among note intervals and durations: notes which have the higher weights are selected as boundaries in the melodic surface and musical phrases are hence highlighted. A quantization of musical phrases in four different levels, regarding both intervals and durations, is then performed. This operation can be seen as the analogous of stemming for textual documents: quantization partially overcomes the problem of variants in musical phrases, that is the presence of phrases with the same structure, but with small changes in durations and interval structure. Musical phrases, with their different level of quantization, can then be indexed using standard techniques for textual IR. This methodology was tested on a experimental collection of tonal western music; results are presented in (Melucci, Orio 1999).

Musical phrases can also be used as content descriptors like index terms of textual documents. Hence the method for automatic melodic segmentation can be combined with automatic construction of hypertext techniques (Agosti et al. 1997), obtaining a network of musical documents that the user can navigate. After having queried the system, the user is provided with a number of musical documents from which s/he can initiate browsing. Automatically generated hypermusic provides different kind of links, which permits: to retrieve documents being similar to those retrieved, but unretrieved because they were judged unrelevant by the system having a low score; to look up content describing phrases, especially the phrases absent in the user's query; to retrieve documents indexed by a particular phrase; to reformulate the query whenever the user wants to use similar phrases to retrieve previously unretrieved documents.

Author Information

Massimo Melucci
Department of Electronic and Computer Engineering
University of Padova
Via Gradenigo, 6/A
35131 Padova
Phone: (+39) 049-8277694. Fax: (+39) 049-8277699
melucci@dei.unipd.it

Nicola Orio
Department of Electronic and Computer Engineering
University of Padova
Via Gradenigo, 6/A
35131 Padova
Phone: (+39) 049-8277665. Fax: (+39) 049-8277699
orio@dei.unipd.it

Suggested Readings

Cambouropoulus, Emilio. 1997. Musical rhythm: a formal model for determining local boundaries. In M. Leman, editor, *Music, Gestalt and Computing*, Springer-Verlag, Berlin, pages 277-293.

Melucci, Massimo; Orio, Nicola. 1999. Musical information retrieval using melodic surface. In *Proceedings of ACM Digital Libraries Conference*, Berkeley, CA. In press.

Agosti, Maristella; Crestani, Fabio; Melucci, Massimo. 1997. On the use of information retrieval techniques for the automatic construction of hypertexts. *Information Processing and Management*, 33(2): 133-144.

Representing Musical Information for Retrieval

Abstract

Dozens, perhaps hundreds, of way of representing music have been developed. Codes for representing music have tended to concentrate on selective understandings of what music is. To the engineering community, music is sound and frequency is an accurate absolute measure. To singing teachers, the practice of solfegge instills a valuable generalized notation of scale degree. This is a relative measure; indeed generalization and relativity are synonymous. In specific application domains—sound, notation, academic analysis, and "bibliographic" melodic searching—a great range of codes is currently in use. All of these codes have some kinds of information in common. However, the elements unique to each domain are somewhat problematical for the notion of a universal standard for the interchange of musical data, which has tended to be the Holy Grail in discussions about music applications.

The absence of an interchange code is sometimes taken to reflect the absence of a comprehensive code for representing music, but several which are quite comprehensive exist. Thus when novices proceed to develop a new code, they simply enlarge the pluralism of the problem that their successors will face. The biggest obstacle to the development of an interchange language may be the idealism with which the challenge is approached. An interchange language that will satisfy the needs of scholars, performers, teachers, and rank-and-file musicians as well as the engineers, hardware designers, and software engineers is a lofty goal and possibly less desirable than was once thought.

Interchange codes that operate within specific domains (e.g., sound or notation), on the other hand, are practical and in some are already operational. When we confront such applications as melodic searching, we notice a new sphere of needs related to algorithmic suitability, efficiency, and reliability. Many human factors—cognitive, perceptive, and cultural—become relevant here. A paucity of programs (reflecting a paucity of data) has deprived us of opportunities for testing, comparison, and validation. These circumstances are changing rapidly, as the critical need for larger data pools is being met by collaborative projects. These include the databases of the Center for Computer Assisted Research in the Humanities and their translations for use in MIDI applications, printing applications, and the analysis toolkit *Humdrum*. Similarly, the Web tool *Themefinder*, based on software by David Huron (with interfaces and extensions by Andreas Kornstaedt, Craig Sapp, and others), facilitates melodic searching of more than 10,000 thematic incipits of classical music. Those who best understand the divergence of needs in existing domains of music applications will be best equipped to tackle the next level of the problem.

Author Information

Eleanor Selfridge-Field
Center for Computer Assisted Research in the Humanities
Braun #129, Stanford University
Stanford, CA 94305-3076
Tel. (650) 725-9242; Fax (650) 725-9290
esf@ccrma.stanford.edu
http://www-ccrma.stanford.edu/~esf

Suggested Readings and Web Sites

Selfridge-Field, Eleanor (ed.). 1997.* Beyond MIDI: The Handbook of Musical Codes. *Cambridge, MA: The MIT Press.

Hewlett, Walter B., and Eleanor Selfridge-Field (eds.). Melodic Similarity: Concepts, Procedures, and Applications (Computing in Musicology, 11). Cambridge, MA: The MIT Press.

Full-score encodings at CCARH: http://musedata.stanford.edu/databases/index.html

Themefinder Project for searching melodic incipits: http://musedata.stanford.edu/databases/themefinder/index.html

Searching for Sound and Music at IRCAM

Abstract

IRCAM is a non-profit organization focusing on research in music-related technologies, music composition and concert production. It is also a degree-granting higher education institution in computer music and musicology.

This talk will briefly present an overview of current activities related, or applicable, to various aspects of music information retrieval (e.g., off-line indexing of audio material), such as :

- Implemented systems :
  - Extended metadata for scores and recordings (search by instrumentation)
  - Perceptual search for sounds in a large database (search by perceptual characteristics)

- Research :
  - MPEG 7 proposal for extended audio description (metadata)
  - Fundamental frequency recognition (melody extraction)
  - Incremental music material parsing into motifs
  - Feature extraction and temporal segmentation of audio signals (automatic identification)
  - Recognition of audio excerpts in a large database (search for specific recording)
  - Musical similarity perception (search for similar recording)

Author Information

Michael Fingerhut
Director, Multimedia Library
IRCAM – Centre Georges-Pompidou
1, place Igor-Stravinsky
75004 Paris, France
Office : (+33) (0) 1 44 78 48 53 Fax : (+33) (0) 1 44 78 15 40
mf@ircam.fr

Suggested Reading

Assayag, Gérard *et al.* : Guessing the Composer's Mind : Applying Universal Prediction to Musical Sequences. *Proceedings of the International Computer Music Conference* 1999 (to appear).

Doval, Boris *et al.* : 1993. Fundamental Frequency Estimation and Tracking using Maximum Likelihood Harmonic Matching and HMMS. *Proc. IEEE ICASSP* 1993 : 221-224.

Dubnov, Shlomo *et al.* : Universal Classification Applied to Musical Sequences. *Proceedings of the International Computer Music Conference* 1998.

Fingerhut, Michael : 1999. The IRCAM Multimedia Library : a Digital Music Library. *Proceedings of the IEEE Forum on Research and Technology Advances in Digital Libraries* 1999 : <URL:http://mediatheque.ircam.fr/-articles/textes/Fingerhut99a/>

McAdams, Stephen *et al.* : 1999. The Perception of Musical Similarity. *Proceedings of the International Symposium on the Neurosciences of Music* 1999: in preparation.

Misdariis, Nicolas *et al.* : Validation of a Multidimensional Distance Model for Perceptual Dissimilarities among Musical Timbres. *Proceedings of the Joint International Congress on Acoustics and Meeting of the Acoustical Society of America* 1998 : <URL:http://mediatheque.ircam.fr/articles/textes/Misdariis98a/>

Rossignol, Stéphane *et al.* : 1998. Feature Extraction and Temporal Segmentation of Acoustic Signals. *Proceedings of the International Computer Music Conference* 1998 : <URL:http://mediatheque.-ircam.fr/articles/textes/Rossignol98a/>

Worms, Laurent : 1999. Reconnaissance d'extraits sonores dans une large base de données. [Sound excerpts recognition in a large database]. Report of an internship at IRCAM (restricted access). *Ecole nationale supérieure de physique de Marseille* 1999.

CUIDAD Working Group Web site : <URL:http://www.ircam.fr/cuidad/>

Studio OnLine Web site : <URL:http://www.ircam.fr/sol/>

The transparencies for this talk : URL:http://mediatheque .ircam.fr/articles/textes/Fingerhut99c/

Trends in Musical Iconography: Concerning Words (Word Soup), Numbers (Iconclass), and Picture Retrievals in a Polyglot World

Abstract

Fundamental to the field of musical iconography is the researcher's ability to retrieve musically relevant images from large art "imagebases" of a general, rather than a subject-specific, nature. Most of the great picture libraries and print rooms in the world, whether in the US (like that of the Fogg museum, or the National Gallery of Art), or abroad (like that of the Louvre, or the Rijksmuseum) have classed their images for decades by artist, by period, and/or by nationality/school--aspects which are significant to art historians, less so for those seeking visual evidence on specific subjects. Since WWII, a small and growing number of libraries and museums have been experimenting with more or less detailed, subject-oriented indexing of their art collections. In recent years, some have even made great strides in mounting their indexed images on the Web.

Currently there are two contrasting methods of image indexing and retrieval in general use, namely natural-language (word-based) systems, and symbolic or alphanumeric ones. The most popular of the latter type is the Dutch Iconclass system (http://iconclass.let.ruu.nl/). I will attempt to review the perceived advantages and drawbacks of both these methods, with special attention to a new (purists might call it a "wild west") development in computerized, word-based retrieval, namely "Word Soup." Currently it is up and running at the Fine Arts Museums of San Francisco (FAMSF) Web site (http://www.thinker.org/imagebase/index.html). Some reflections on the adequacy of word-based systems to the demands of multilingual retrieval will be offered.

The FAMSF site also demonstrates an exciting new development in the delivery of pictorial detail over the Web, without having to wait forever for images to load. It is a user-friendly zoom technology, developed in collaboration with the Computer Science Division of the University of California at Berkeley, called "GridPix." Will we some day be able to "zoom in" on an image as effectively with words (verbally) as we now can visually with GridPix technology? "Hypericonics" (v. Nauta reference below) may assist us with this challenge.

Author Information

Thomas F. Heck
Professor of Musicology and Head,
Ohio State University Music & Dance Library
Sullivant Hall, 1813 N. High St.
Columbus, OH 43210-1307
(614) 292-2319, Fax: (614) 292-8012, Home: (614) 442-8226
insights@aya.yale.edu

Suggested Readings

Heck, Thomas F. 1999. *Picturing Performance: The Iconography of the Performing Arts in Concept and Practice.* Rochester, NY: Univ. of Rochester Press.(In press, September release.)

Seebass, Tilman, and Tilden Russell, ed. 1984-. *Imago musicae*, Vol. 1-. See the "Bibliographia, 1975-1981" section found in the first volume of this yearbook, and "Bibliographia" updates in every volume since.

Musical iconography: a manual for cataloguing musical subjects in western art before 1800. Ed. by Howard Brown and Joan Lascelle. Cambridge, MA: Harvard University Press, 1972.

Nauta, Gerhard Jan. 1993. HYPERICONICS: Hypertext and the social construction of information about the history of artistic notions. *Knowledge Organization* 20 (1): 35-46.

Seebass, Tilman. 1992. Iconography. In *Ethnomusicology I: An introduction.* London; New York: Macmillan; Norton, 238-44. Prof. Seebass has announced a new Web site (9/1999-?) for music-iconographical news at the University of Innsbruck. See the Ico-News link at: http://www.uibk.ac.at/c/c6/c619/.

Searching a Web-Based Thematic Catalog

Abstract

Thematic catalogs for music are significantly enhanced when there are indices to the themes. For the most part, such thematic indices are a product of the last forty years with the advent of the computer and its ability to order large masses of data. Most indices rely on alpha-numeric music representations, and an index arrangement by interval, key, or scale step. Most thematic catalogs include only the beginning of a musical piece or movement; or the incipit. As a consequence, thematic indices are left-anchored arrangements of data, exactly like a dictionary of words. I am reporting on an examination of search algorithms for musical incipits, and the development of a user interface.

In October 1994, I mounted on the web a thematic catalog of the first book of viola da gamba pieces by Marin Marais (http://orpheus.ucsd.edu/music/ghb.html). Each of the 112 musical incipits is presented as a gif file on its own html page, with links established between registers of printed works, manuscripts, recordings, and title index. The initial goal of the experimental thematic catalog was to make it simple enough that any musician could use it without having to learn a special system. This goal was accomplished by displaying the incipits as an image in musical notation.

Another goal for the experiment was to develop a method of searching for specific incipits that would seem natural to most musicians. This goal was seen to be difficult for two reasons; there was no standard for representing and searching music in the computer, and there was no effective musical user interface. For this project, intervals were chosen as the searching element, and a search will find transpositions of an incipit.

A number of user interfaces for searching the thematic catalog were examined and rejected. However, the common characteristic of literate musicians is the ability to read music. Therefore in April 1998, I decided to use the javascript language along with the graphic capabilities of the web for the user interface. It consists of an image map from which the user selects a sequence of notes. As the database is relatively small, only two or three notes are needed to retrieve a reasonable list of results.

This user interface is relatively crude, and needs to be refined. While the search process described here is limited to thematic catalogs, it could be expanded to search for true themes and motives much as do most word processors for partial strings of text. There is, however, much work to be done to develop such a system. For example, one must examine the methods needed to reduce the number of false hits; the role of rhythm and meter; and how to create a fuzzy search which will return all transpositions and rhythmically altered versions without missing important versions of a theme.

Author Information

Garrett H. Bowles
Music Library
University of California, San Diego
La Jolla, CA 92093
Office: (858) 534-1267 Fax: (858) 534-0189
gbowles@ucsd.edu

Suggested Readings

Bowles, Garrett H. 1978. *The Computer-Produced Thematic Catalog; An Index to the "Pieces de Violes" of Marin Marais*. Thesis (Ph.D.): Stanford University.

Brook, Barry and Richard Viano. 1997. *Thematic Catalogues in Music; An Annotated Bibliogrphy*. 2d ed. Stuyvesant, NY: Pendragon Press.

Hinkle, Win and the Bill Evans Jazz Resource. 1999. *Bill Evans Thematic Catalog* (http://www.billevansjazz.com/34skid/html/body_thematic_catalogue.htm)

RISM. 1996. *RISM Music Manuscripts after 1600*Indexes 285,000 bibliographic records and a thematic catalog (with graphical images of music incipits) for music manuscripts written from around 1601 to 1850 (http://www.rism.harvard.edu/rism/DB.html)

Roland, Perry. 1999. *The "Preposterous Muddle" Revisited: An XML Thematic Catalog DTD*. (A paper from the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing conference, June 9-13 at the University of Virginia in Charlottesville, Virginia.) Proposes a DTD for thematic catalogs. (http://www.iath.virginia.edu/ach-allc.99/proceedings/roland.html)

University of North Texas Music Library. 1995. *Lully Web Project* A pilot project multimedia catalog of the collection. (http://www.unt.edu/lully/welframe.html)

The Humdrum Toolkit: Software Tools for Music Research

David Huron

Humdrum provides a general-purpose set of software tools that facilitates research-related music applications. Humdrum is able to represent a large range of music information, from French lute tablatures to Dagomba dance gestures. In addition, the toolkit provides software for searching, extracting, classifying, labeling, comparing, and otherwise manipulating the musical data. The Humdrum tools are increasingly used to build more specific applications software. The presentation will include a demonstration of "Themefinder" -- a web-based "name-that-tune" search engine created using Humdrum tools.

Music Processing Using Colour

Abstract

From interviews and an Internet survey, we have observed that a significant number of composers and arrangers prefer to use pen and paper to write music, even though there is a large range of software now available for writing and arranging. Therefore, it would be very helpful if some image processing methods could be applied directly to manuscript music, to enable musicians to work in their preferred medium of paper, but gain some of the benefit of computer assistance. This could make tasks such as arranging and part writing more efficient.

Due to the superimposition of notes and symbols on the staves, and the wide variety of characteristics of musical notations (e.g. size and shape), image processing of music has been a challenge for researchers. Most projects have focused on processing bi-level printed musical scores which had white as the background colour and black for both staves and musical symbols. In this case, it is difficult to separate the staves and other overlapping musical symbols because they are all black. However, since the technology of colour inkjet printers and colour scanners has been significantly improved, and the huge drop in price makes them very affordable, we propose to use different colours for staves and handwritten musical symbols. By setting the colour of the staves to a particular colour value, we expect it should be easier to remove those pixels to enable recognition of the musical symbols alone.

A variety of coloured staves were printed so that we could experiment with using different coloured pens and pencils. We discovered that when a dark colour is put on a light colour, the overlapping area is shown as a dark colour and the light colour almost disappears. Black (including lead pencil) is especially strong and covers other colours. In processing music, it is more important to recognise the musical symbols rather than the staves, therefore, the best format is to have light colours for staves and dark colours (e.g. black or gray) for musical symbols.

Scanned images are originally stored as an RGB (Red, Green and Blue) format. When musical scores are scanned, the original RGB values of both stave and symbol are not reproduced accurately and can vary on the same page. Even the white background does not stay as white as the original. Therefore, we need to develop a method of effectively and accurately separating the background, staves and musical symbols from a coloured musical score by simply using the colour information we can obtain from the original RGB values.

There are other colour spaces to be considered, such as HSV (Hue, Saturation and Value), CMYK (Cyan, Magenta, Yellow and Black), etc. By transforming RGB to other colour spaces, we have found that better information is available to distinguish different colour groups in the scores. For example, a high saturation value indicates coloured staves, and the black of a CMYK space can be used to identify pencil writing. We will apply these observations to a new recognition technique.

Author Information

Fang-Yi Lin
Department of Computer Science
University of Canterbury
Private Bag 4800, Christchurch, New Zealand

karen@psyc.canterbury.ac.nz

Tim Bell
Department of Computer Science
University of Canterbury
Private Bag 4800, Christchurch, New Zealand
tim@cosc.canterbury.ac.nz

Suggested Readings

Bainbridge, David. 1997. Extensible Optical Music Recognition. *Ph.D. Thesis, University of Canterbury, New Zealand.*

Lin, Fang-Yi and Bell, Tim. 1999. Music Processing Using Colour. *IVCNZ'99 Conference Proceedings (will be published in September).*