A Framework for the Automatic Discovery and Description of Musical Structure

Elaine Smyth *Kevin Curran* *Paul McKevitt* *Tom Lunney*

School of Computing & Intelligent Systems, University of Ulster, Magee Campus,
Northern Ireland, BT48 7JL, UK

ABSTRACT

Automated discovery of recurrent patterns in music plays an important role in computational music analysis. Patterns emerge from repetitions within the music itself, from which musical structure can be inferred. Many papers have been written in relation to automated musical analysis resulting in a number of different approaches to structural analysis and representation. This paper aims to provide an overview summarising some of these approaches, from matrix-based systems to XML. MPEG-7 metadata standard will also be introduced and the proposed architecture of the AMUSED project for automated MPEG-7 structural analysis and description will be presented.

1. INTRODUCTION

The automated discovery of recurrent patterns in music plays an important role in computational music analysis. Patterns emerge from repetitions within the music itself, and it is from these patterns and repetitions that structure can be inferred. Structure is a defining element within music, whether vertically in chords and harmonies, or horizontally in measures and phrases, without it music would simply be noise with no direction or purpose. Every musical piece contains some element of similarity, whether at verse and chorus level, or at note pattern level. The identification of structure is useful within a number of application areas, such as Query-By-Example (QBE), audio-fingerprinting and identification, indexing and archiving, search and retrieval, audio classification and analysis, to name a few.

The human ear can easily identify similar sections of music, even a musically untrained person can usually pick out high level structures like verse and chorus. However, to ask a computer to do the same is no simple task. Complex matching algorithms must be employed to identify patterns within a piece; these patterns are then used as a basis for comparison to identify repetition; eventually a structural 'understanding' of a piece is formed and a description is provided. Many attempts have been made to replicate the human interpretation of structure within music, with varying levels of success. However, very little has been done with respect to utilising the relatively new MPEG-7 meta-data standard to model and represent musical structure. The MPEG-7 standard permits the description of the hierarchical nature of music at varying levels of detail; from the power of the actual audio signal itself to the high level description of song form. The resulting MPEG-7 compliant descriptions could then be used as the basis for further processing by other systems which seek to utilise this type of structural information, such as visual based music applications or error concealment modules within streaming media algorithms, among others; made possible through standardisation.

This paper provides an overview of a selection of previous research carried out in relation to structural musical analysis. Following this, XML's application to music will be discussed which, in turn, leads us into the XML-based MPEG-7 Audio Standard (15938-4). To end the paper, the AMUSED system will be introduced and overviewed and a description of the proposed system architecture will be provided.

2. AUTOMATED MUSIC STRUCTURE ANALYSIS

Just as a paragraph, in written and spoken language, can be broken down into sentences, sentences into words, words into syllables, etc, music can also be broken down in a similar manner into its constituent parts. West, et al, (1991) provide examples of eight such constituent parts in increasing complexity: note, duplet, triplet, bar, motif, phrase, stanza, and movement. Each of these can be arranged hierarchically into a structure which is representative of the musical piece they comprise. For example, notes can be grouped together to form bars, which can be grouped together to form phrases, and so on, until the highest possible representation is reached. However, there is more to music than the simple organisation of notes into bars; chords, harmonies, melody, rhythm, and timbre, etc, are all important elements within the musical mix.

It is almost universally agreed that patterns and repetition are also defining characteristics within music; the discovery of such patterns and repetitions are important indicators to musical structure. Patterns emerge from repetitions within the music; repetitions can range in size from simple repeating note sequences, to recurring sections or phrases within a musical piece. At a high level this would include the verse, chorus, verse, chorus, structure. The automated discovery of recurrent patterns within music is an important problem within computational structural music analysis. Music has a rich internal structure consisting of multiple types of substructures, and, as mentioned, analysis of this structure can be useful to a variety of applications.

There have been many papers written in relation to automated musical analysis and to outline them all here would be a lengthy process. Techniques developed to discover structures within music have resulted in a number of different approaches. This overview will attempt to give a representative summary of some those differing approaches.

Heinrich Schenker is perhaps the grandfather of structural analysis techniques, and is almost certainly the most cited, in particular his Fundamental Structure theory for tonal music. In brief, his Fundamental Structure is made up from three 'layers': Background, Middleground and Foreground. These layers are based on diminutions (of which there are 4 types); if layers of music are stripped away from the surface diminutions which span larger and larger sections of a piece below the surface can be identified. He argued that the Foreground of music was built up from simple diminutions in the Background; diminutions have notes added to them in successive layers from the Background forwards. Schenker used traditional musical notation with modified implications using his own symbols. (Pankhurst, 2005).

Another commonly cited theory is Lerdahl and Jackendoff's Generative Theory of Tonal Music (GTTM). Based on a linguistic theory (Chomskian) which takes in Schenkerian analysis, it attempts to describe how a listener perceives structure within music; particularly from the perspective of an experienced listener. They define four hierarchical dimensions: Grouping Structure, Metrical Structure, Time-Span Reduction, and, Prolongational Reduction. These dimensions represent different aspects of structure within a piece. In addition to these dimensions, they also define two sets of rules which are designed to lead one to arrive at any one of the four dimensions: Well-Formedness rules and Preference rules. The rules of Well-Formedness are not unlike the grammatical rules of language; they define what a well-formed, or syntactically correct, structure is. The Preference rules simply allow preference or importance to be given to certain structures within levels; they help select one possibility from a number of possibilities discovered by the well-formedness rules (Valkonen, 2000).

More recent approaches show ever more varied and complex techniques for discovery of patterns and musical structure, in varying levels of detail. David Temperley, in collaboration with Daniel Sleator, present a computational theory of music cognition; Temperley's work, upon which the computerised approach is based, is heavily influenced by Lerdahl and Jackendoff's GTTM, mentioned above. Like Lerdahl and Jackendoff, music cognition is explained by means of a system that generates structural descriptions based on a set of preference rule systems. These preference rule systems consist of well-formedness and preference rules of the type used in GTTM. Each takes a representation (pitch, expressed in MIDI numbers, onset and offset times of notes) of a musical passage as input and generates a structure as output that is intended to describe certain aspects of how the musical passage is interpreted by an expert listener. Temperley defines six aspects of musical structure: metrical, phrasing, contrapuntal, pitch-spelling, harmonic, and key (Meredith, 2002).

Melisma Music Analyser is the computer implementation of Temperley's theory. It consists of several modules which analyse musical input to produce structures for each of the structural aspects mentioned above. Output from the system is comprised of various alphanumeric figures and symbols used to represent different information. For example, output from the Harmonic Analyser module consists of a time-point in milliseconds representing the start of a segment; a metrical strength indicated by a series of X's; the result of the harmonic analysis is the root of the segment; followed by a graphical representation of the root's position on the 'line-of-fifths'. As pitch-spelling is combined with harmonic analysis, the Tonal Pitch Class (TPC) is labelled together with its 'line-of-fifths' representation (Sleator and Temperley).

Foote and Cooper (2001), (2003), base their novel approach on the use of a two-dimensional matrix representation of musical information. An audio file is visualised as a square; time runs from left to right, as well as from top to bottom; the bottom left-hand corner of the image represents the start of the piece, while the top right corresponds to the end. Similarity is signified by light (similar) and dark (dissimilar) cells within the matrix. Repetitive similarities show up as a checkerboard pattern and repeating themes are visible as lines parallel to the diagonal. Coarse structure can be inferred from the checkerboard patterns within the matrix. Segments can be discerned from square regions of high similarity along the main diagonal allowing the identification of Intro, Verse, and Chorus blocks.

Lu, et al, (2004) present a similar matrix-style methodology. They offer improved feature extraction through a Constant Q Transform and a novel similarity measure between features. From the resultant self-similarity matrix, an adaptive method is used to identify all significant repeating patterns; and from this, heuristics are used to analyse the repeating patterns further. Each repeating segment is labelled to obtain musical structure, based on standard musical form notation (ABAA, for example). The ultimate result from this analysis is a fairly course description of musical structure, or song form.

Dannenberg (2002) uses repetition as a basis for structural understanding of music. The system takes actual digital audio as input and produces a simple 'explanation' of the music in terms of repetitions and implied structure as output. The structural explanation is non-hierarchical; output form the system is a string of symbols, e.g. AABA, in the usual musical form notation. Vercoe and Chai (2003) present a similar system which automatically identifies repetition and labels high level structure within music; as do, Maddage, et al (2004); a non-hierarchical coarse description of a song is ultimately produced indicating Intro, Verse, Chorus, etc.

Weyde (2001) bases his approach on discovering rhythmic structure within music. Information at motif and phrase level is extracted; low level data is deemed un-useful as it does not contain enough structural information. The basic idea is to determine groupings for sequences of notes and similarity relations of groups. The end product of the system is a series of horizontal groupings with defined relations. A further paper by Weyde and Wissman (2004) uses musical maps to visualise structure graphically; the relations within music itself are represented as nodes and 'edges'. Various musical objects are defined; these can range in size from a note to an entire collection, can be more abstract concepts like harmonic function, or parts of musical pieces like a motif or chord. Relations can be even more varied; containment, similarity, succession, rhythmic and harmonic relations, for example. The maps adequately show the hierarchical relationships between musical elements.

A lot of the previous research outlined here produces a relatively high level, coarse, non-hierarchical description of musical structure. These relatively non-hierarchical representations should really be referred to as musical song form, as this is what they ultimately resemble. Temperley and Sleator's system is perhaps the most comprehensive; it models musical structure from six different perspectives. However, it is somewhat limited in the inputs it can handle, ('notelist' – pitch (midi#), onset and offset times), and the representation it produces: alphanumeric figures and symbols. The musical maps, produced by Weyde and Wissman (2004), also show a reasonable degree of structural information and relationships, but the system is a self-contained object oriented environment for musical visualisation; information produced by the system is for use purely within the system's scope, there is no output other than the visual representation on screen. It would be extremely valuable to have a system which analyses and represents musical structure in a standardised way; the output of which would be useful to any number of applications, even those outside music theory.

Wiggins, et al, (1989) go some way towards providing a standard representation; they apply an abstract representation to model music and its inherent structure; known as CHARM (Common Hierarchical Abstract Representation for Music). Although this is a system of representation and does not perform analysis, it is mentioned here because it provides the means with which to model discovered musical structure. Within the system notes correspond to 'events'; events are made up from a set of tuples consisting of: {id, pitch, time, duration, timbre}. These events can be grouped to form 'constituents'; constituents are a higher level description for a collection of events from which a hierarchy can be built. For example, an event can appear inside a constituent, this constituent may then appear inside a further, higher level constituent, and so forth, to create a hierarchical abstract representation of a musical piece. The CHARM system may produce a somewhat standardised representation, but it is not on par with the complex representation possible with XML-based notation; its main purpose, perhaps, to satisfy a computing requirement. It is also unclear how easy it would be to model differing structural aspects based on the defined event/constituent groupings; however, this would be dependant on particular implementations of the representation. In addition, XML is already standardised and can be easily parsed for specific information because it is text-based.

3. XML AND MUSIC NOTATION

XML (eXtensible Markup Language) is a standardised text-based method for describing the structure of data, readable by both humans and machines. It provides a very rich schema for defining complex documents and data structures, which is also extendable. Text-based representations are ideal for music, especially when these representations can be automatically extracted from sound files. Music does not lend itself to easy interpretation by computer, unlike language which can be easily represented and understood by computers because of the ANSI standard. Of particular benefit to music applications is the fact that XML allows support for multiple descriptions of the same data, and hierarchies are as fundamental to XML as they are to music notation. XML's ability to represent both hierarchies and the relationships between hierarchical elements is beneficial. What makes XML more attractive is the fact that it is platform independent, and standard tools can be used to validate code (Castan, 2001).

There are a variety of XML-based markup projects in existence that are specifically aimed at music. XScore, the Music Encoding Initiative (MEI), MusicXML, and Music Content Markup Language (MCML), to name a few. The reader is referred to http://xml.coverpages.org/xmlMusic.htm for a fairly comprehensive list of existing markup initiatives. Although all are based on XML, they do have varying purposes. For example, XScore is an application of XML for describing the musical score (Grigaitis, 1998), and MCML was developed as part of another project (MIDILib) for content-based queries and navigation (Schimmelpfenning, et al, 2000). While MEI strives to meet a broader range of music applications, it also avoids the confusion of other XML standards by using familiar names for elements and attributes, for example, <note> and <chord> (Roland, et al, 2002). MusicXML is perhaps the most mature endeavour for music encoding using XML. It has its roots in academia and has made its way into a number of commercial applications and has even been proposed to OASIS (Organisation for the Advancement of Structured Information Standards) for consideration to become a recognized standard (Stewart, 2003).

There are a few tools which harness the power of XML-based markup languages; as mentioned, the MIDILib project utilizes MCML; in addition, MiXA, VExPat and Sharpeye, among others, make use of MusicXML. MiXA is a web based musical annotation system (Kaji, et al, 2004), whereas VExPat is pattern extraction system which uses the MusicXML representation of music as a basis for analysis, score rendering and MIDI execution (Santana, et al, 2003). Sharpeye is a music reader which converts a scanned image of printed music into a MIDI, NIFF or MusicXML file (Jones, 2005). There do not appear to be any XML applications which deal specifically with the issue of identification and representation of structure within music, in the context of the current project. However, MUSITECH, which is an object-based model of music, can transform its Java-based objects into XML format for storage and data exchange. This system concentrates on representing layers of music from different aspects, defined by abstraction, aggregation and structure, making it possible to create different views of a piece of music (Weyde, 2004). Although, it is unclear what exactly is captured in the XML representations, or whether musical structure is adequately modelled.

A potential drawback is that most current DTD's restrict themselves to Common Western Music Notation (CWMN or CMN), with some including tablature. MusicXML, for example, was designed to represent musical scores and sheet music, specifically common western musical notation from the 17th century onwards (Stewart, 2003). MPEG-7 is a further XML-based standard, but is more general in its approach; it is not based on the representation of sheet music or scores, but can deal with the representation of music directly from the audio file; in addition it can also deal with spoken content. Analysis of the sound file frees the representation somewhat from the restrictions of a specific notational approach, like CWMN. The next section provides a very brief overview of MPEG-7 Audio (Part 4).

4. OVERVIEW OF MPEG-7 AUDIO

The MPEG-7 standard currently consists of 8 parts, part 4 of which deals specifically with the description of audio data; formally recognised as 15938-4. The main MPEG-7 elements are Descriptors (D), Description Schemes (DS), and a Description Definition Language (DDL). Ds are intended to describe low-level audio features; they are the building blocks of the system. DSs are designed to describe higher level audiovisual features. DSs produce more complex descriptions by combining multiple Ds and DSs and declaring relationships among the description components. DDL provides the descriptive foundations through which users can create their own Ds and DSs (Salembier, et al, 2001).

The Audio Framework tools are applicable to the description of general audio; a graphical representation of the Framework is provided in Figure 1. The generic Audio Framework contains low-level descriptors designed to provide a basis for the construction of higher level audio description schemes. The Low-Level Descriptors (LLDs) permit the description of an audio signal's spectral and temporal features. In the first version of the standard there were seventeen LLDs for general use in a variety of applications. Since 2004, there has been an extension to the original standard (Amendment 1) to include additional Ds for such things as background noise level, balance, bandwidth, etc, with proposals for a second amendment to include Ds to describe audio intensity, rhythmic patterns, chord patterns, and so forth (ISO/IEC, 2002) (Gruhne, 2004).

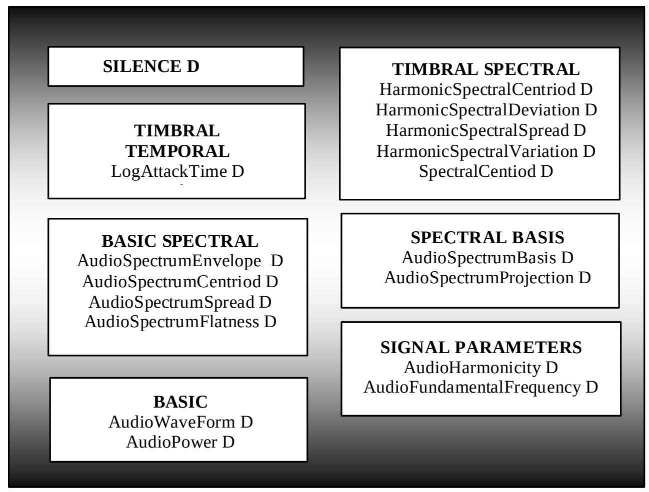

Figure 1 Audio Framework

In addition to the LLDs there are five general sets of high-level audio description tools, which aim to encompass some application areas, such as sound recognition, musical instrument timbre, spoken content, melodic contour and melody. These specialised tools may be used in conjunction with the other tools within the standard. The high-level tools provide both functionality and also serve as examples of how to use the low-level framework (ISO/IEC, 2001).

MPEG-7 Ds and DSs have been successfully utilised by a number of projects over the last few years in a variety of application areas (Batke, et al, 2004), (Celma, et al, 2004), (Dumouchel, et al, 2003), (Eisenberg, et al, 2004). These can be broadly categorised into one of the following areas: Query-By-Example (QBE), fingerprinting for audio identification, indexing and archiving, MPEG-7 authoring tools, audio analysis, audio classification, and Digital Rights Management (DRM). Musicstructure.com is perhaps the only existing MPEG-7 based application which deals solely with the analysis and representation of musical structure, (Casey, 2004), but again this system only performs partitioning into a relatively high level structure.

MPEG-7 holds great promise in relation to the useful description of audio. It will have many applications, both within music theory and within information search and retrieval, as well as many as yet unknown application areas.

5. AMUSED SYSTEM ARCHITECTURE

There are already quite a few applications in existence which utilise the power of MPEG-7 Audio; ranging from applications which automatically extract low-level descriptions, (Crysandt, 2003), to applications which extract and use descriptions to visualise the structure of music, not unlike the matrices discuss earlier in this paper (Casey, 2004). Other systems developed have been mainly based on the query-by-example paradigm or audio fingerprinting systems for indexing and archiving; there are very few pure musical structure analysis tools. The aim of the current AMUSED project is to develop a Java-based real-time architecture which will take digital musical audio (in most common formats) as input and produce MPEG-7 compliant descriptions of structure as output. A diagram representing an overview of the proposed system architecture is shown in Figure 2.

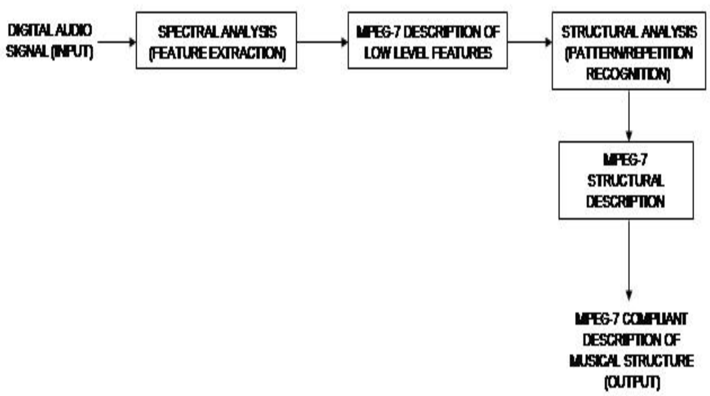

Figure 2 Overview of AMUSED System Architecture

The aim of this system is to achieve real-time analysis of digital audio and subsequent creation of MPEG-7 compliant structural descriptions. A digital audio signal is received into the system as a data stream. Spectral analysis is performed on the incoming data stream to accumulate low level information about the audio signal; signal power, spectrum spread, fundamental frequency, etc. This information is ‘extracted’ from the signal and described using the MPEG-7 Audio LLDs. The structural analysis module takes the output from this module and performs further analysis to identify patterns, repetition and regions of self-similarity. This information is then used as a basis for the construction of a hierarchical model of musical structure which is described by the MPEG-7 structural description module. The final output from the system is in the form of an MPEG-7 compliant XML file, which can be used within other systems.

6. SUMMARY

This paper provided a review of previous research and an overview of the MPEG-7 standard, together with a synopsis of the proposed system architecture. Most of the systems reviewed provide high level song form descriptions with no real appreciation of the structure that lies beneath. Those which did model structure at a more in-depth level lacked any standardised output, limiting usefulness outside their own project scope. The use of the XML standard for music notation is a step in the right direction; however, some of the projects outlined here were rather simplistic in nature and didn’t capture anything that useful about the music.

Built on top of the XML base is the MPEG-7 standard. MPEG-7 offers a standardised scheme for the description of audio, among other media. MPEG-7 contains specific extensions to facilitate the accurate description of digital audio content. Some applications are already in existence which utilise MPEG-7; however, very few offer the automated description of musical audio structure, indeed, most are QBE-style applications. The AMUSED system aims to overcome the inadequacies identified and provide a novel system for the automated extraction, analysis and description of digital audio structural characteristics and produce standardised, MPEG-7 output.

References

Batke, Jan-Mark, Eisenberg, Gunnar, Weishaupt, Philipp, Sikora, Thomas, (2004) *A Query By Humming System using MPEG-7 Descriptors*, AES Convention Paper 6137, AES 116th Convention, 8th-11th May 2004, Berlin, Germany

Casey, M, (2004), Musicstructure.com, http://www.musicstructure.com/intro.html

Castan, G., Good, M., and Roland, P. (2001). *Extensible Markup Language (XML) for Music Applications: An Introduction.* Computing in Musicology, Vol. 12, No. 1, pp: 12-20

Celma, Oscar, Gomez, Emilia, Janer, Jordi, Gouyon, Fabien, Herrera, Perfecto, Garcia, David, (2004), *Tools for content-based retrieval of audio using MPEG-7: the SPOffline and the MD Tools*, AES 25th International Conference, 17th-19th June 2004, London

Cooper, M. and Foote, J. (2003) *Summarizing Popular Music via Structural Similarity Analysis*. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. Mohonk Mountain House. New Paltz, New York October 19-22, 2003

Crysandt, Holger, Tummarello, Giovanni, Piazza, Francesco, (2004), *Java MPEG-7 Audio Encoder*, MusicNetwork 3rd Open Workshop, 13-14th March 2004, Munich, Dannenberg, Roger, (2002), *Listening to Naima: An Automated Structural Analysis of Music from Recorded Audio*, In Proceedings of the International Computer Music Conference, San Francisco, International Computer Music Association, (2002), pp 28-34

Dumouchel, Pierre, (2003), *MPEG-7 Audiovisual Document Indexing System (MADIS*), Testbed Development, RISQ 2003, 14th Edition, Annual Event of the Network of Scientific Information of Quebec (RISQ), Annual Workshop of CANARIE, Montreal, Canada, 20th-23rd October 2003

Eisenberg, Gunnar, Batke, Jan-Mark, Sikora, Thomas, (2004), *BeatBank – An MPEG-7 Compliant Query by Tapping System*, AES Convention Paper 6136, AES 116th Convention, 8th-11th May 2004, Berlin, Germany

Foote, J. and Cooper, M. (2001) *Visualizing Musical Structure and Rhythm via Self-Similarity*. International Conference Computer Music, Habana, Cuba, pp:140-148, September 2001

Good, M. (2001). MusicXML for Notation and Analysis. In The Virtual Score, ed. W. B. Hewlett and E. Selfridge-Field (Cambridge, MA, 2001), MIT Press, 113-124

Grigaitis, R. (1998). eXtensible Score Language (XScore) 0.01. http://grigaitis.net/xscore

Gruhne, M. (2004). *MPEG-7 Audio.* Workshop Music Network, Barcelona, Spain March 2004

ISO/IEC (2002). *Information Technology – Multimedia Content Description Interface – Part4: Audio*, Amendment1:Audio Extensions, ISO/IEC JTC1/SC29/WG11 N4769, http://www.itscj.ipsj.or.jp/sc29/open/29view/29n4854t.doc

ISO/IEC (2001). *Information Technology – Multimedia Content Description Interface – Part 4: Audio.* ISO-IEC JTC 1/SC 29/WG 11, ISO-IEC FDIS 15938-4:2001(E), http://projekt.rz.tu-ilemnau.de/~kfn/seminar13/MPEG-7-Dokument.pdf.

Jones, Graham, (2005), *SharpEye Music Reader*, http://www.visiv.co.uk/about.htm

Kaji, Katasuhiko, and Nagao, Katashi, (2004), *MiXA: A Musical Annotation System*, Proceedings of 3rd International Semantic Web Conference, Hiroshima, Japan, November 2004.

Lu, L. Wang, M., and Zhang, H. (2004). *Repeating Pattern Discovery and Structure Analysis from Acoustic Music Data*, Proc. of IEEE International Conference on Multimedia and Expo (ICME '04), Taipei, Taiwain, June 27-30, 2004

Maddage, N., Changsheng, K., Mohan, S, and Shao, X. (2004). *Content-based Music Structure Analysis with Applications to Music Semantics Understanding,*, ACM Multimedia, New York, October 2004

Meredith, David, (2002), Review of *David Temperley's The Cognition of Basic Musical Structures,(Cambridge, MA: MIT Press, 2001)*, Musicae Scientiae, 6(2), pp287-302, 2002.

Pankhurst, T. (2003). *Guide to Schenkerian Analysis: Fundamental Structure*, http://www.schenkerguide.com/fundamentalstructure.html

Roland, P. (2002). *The Music Encoding Initiative (MEI)*, Proceedings of 1st International Conference Musical Applications using XML, Milan: State University of Milan, pp 55–59, 2002

Salembier, P., and Smith, J. (2001). *MPEG-7 Multimedia Description Schemes*, IEEE Transactions Circuits and Systems for Video Technology, Vol. 11, No. 6, pp: 65-73, June 2001

Satana, Hugo, Dahia, Erneto Lima, and Ramalho, Geber, (2003), *VExPat: An Analysis Tool for the Discovery of Musical Patterns*, Proceedings of IX Brazilian Symposium on Computer Music, Campinas, SP, August 2003.

Schimmelpfenning, J and Kurth, F, (2000), *MCML: Music Contents Markup Language*, Proceedings of International Symposium of Music Information Retrieval, Plymouth, MA, USA, October 2000.

Sleator, Daniel, and Temperley, David, *The Melisma Music Analyzer,* http://www.link.cs.cmu.edu/music-analysis

Stewart, Darin, (2003), *XML for Music*, http://emusician.com/mag/desktop/emusic_xml-music/

Valkonen, Heikki, (2000), *Lerdahl and Jackendoff Revisited, A Generative Theory of Tonal Music, by Fred Lerdahl and Ray Jackendoff,* http://www.cc.jyu.fi/~heivalko/articles/lehr_jack.htm

Vercoe, B., and Chai, W. (2003). *Structural Analysis of Musical Signals for Indexing and Thumbnailing*, ACM/IEEE Joint Conference on Digital Libraries, New York, pp: 27 – 34, ISBN:0-7695-1939-3

West, Robert, Howell, Peter, and Cross, Ian, (1991), *Musical Structure and Knowledge Representation*, in Representing Musical Structure, Edited by Howell, West and Cross, pp1-30, Academic Press Limited, London, 1991

Weyde, Tillman, (2001), *Grouping, Similarity and the Recognition of Rhythmic Structure*, In Proceedings of the International Computer Music Conference 2001, Cuba, 2001.

Weyde, Tillman, and Wissmann, Jens, (2004), *Visualization of musical structure with maps*, In Proceedings of the Conference on Interdisciplinary Musicology (CIM04), Austria, 15-18 April 2004, Edited by Parncutt, R, Kessler, A, and Zimmer, F.

Weyde, T. (2004). *Symbolic Music Representation in the MUSITECH Project,* 4th Music network Open Workshop 2004, Barcelona, Spain

Wiggins, Geraint, Harris, Mitch, and Smaill, Alan, (1989), *Representing Music for Analysis and Composition*, Proceedings of the 2nd IJCAI AI/Music Workshop, August 1989