Modelling Cognitive and Analytic Musical Structures in the MUSITECH Framework

Tillman Weyde
Department of Computing
City University
London, United Kingdom
t.e.weyde@city.ac.uk

Abstract This article describes the music representation in the MUSITECH infrastructure. This integrated representation provides applications with an object model of musical information that allows the modelling of cognitive and analytic musical structures. Based on the concepts of integrated representation and object independence, the model combines multiple levels of music representation (audio, performance, notation, lyrics, form, harmony, metadata) and allows internal and external links of objects on different levels. Some aspects are covered here in more detail, like the relation of notes in performance and scores, the representation of metre and timing, and annotations and metadata. This object model is part of a software infrastructure that provides applications with numerous components and services, including standard western notation and an XML format.

1 Introduction

Music is a complex domain, and musical knowledge has various types of contents, technical forms, and epistemological qualities. Most currently used representations of music are designed for certain aspects of application or processing, e.g. playback of audio recordings, controlling musical instruments, producing printable scores, or finding music of a certain genre or by a certain artist. For these purposes, there are various file and stream formats, which represent music on different levels of abstraction such as in audio files, performance, notation, and metadata, sometimes allowing additional information on lyrics and harmony.

Yet, for many applications based on musical knowledge – like music production, music tuition, music theory, music psychology, and music information retrieval – there are a number of shortcomings of the available formats. Particularly music theory and psychology, but also music production and tuition, need to define musical structures, i.e. sets of objects or parts of musical materials. These can be for instance the voices, chords and harmonic progressions, melodic motif and phrase structures, or large scale form. These structures need to be represented with various relevant properties, such as harmonic functions, melodic range, chord type, rhythmic quality etc., and their internal and external relations, such as similarity relations, leitmotifs, or relations between performance and notation and general music analysis including links to external information. This kind of structures is largely not present in currently used formats or only in a very limited form. Additional problems occur from the representation being based on particular assumptions or restrictions, like being able to represent a note only within a key and metre, that make it difficult to represent the relationship between notes and metre or key.

This situation makes the development and use of software difficult, leading to commercial music applications being limited in the quality, features, and interoperability they can offer, and music researchers often redundantly developing their own formats and basic tools.

2 Representation of Musical Information in MUSITECH

The MUSITECH representation and software infrastructure were developed in a research project at the University of Osnabrück, supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). In this project, an integrated representation of music has been developed based on the two principles of integrated representation and object independence. The principle of integrated representation means that multiple forms and levels of information are combined and connected in the representation. Therefore provisions are made for the inclusion of different types of media and information, ensuring that all information relevant to a piece of music can be used. The principle of object independence means that object information should be independent of its context. This intends that the interpretation of information in an object, e.g. a note object, should not depend on structures in its context, e.g. metre, key, or previous notes. This is to ensure that objects do not have to be changed when the context is changed, e.g. that the modification of key does not change a note and vice versa. This is especially important when there is more than one structure in parallel, e.g. multiple metres or keys, which is important for modern music and for music software, especially for research purposes.

From a perspective of music analysis and software design, this approach offers maximal flexibility. From a cognitive perspective, the independent representation of notes and structures like metre and key is also adequate, as these structures are not directly part of the performed music. They are inferred by the listener, possibly as implied by musician, but not necessarily so. E.g. a composer usually specifies the metre when she or he writes music down, but often music is not written down and the perceived metre may be ambiguous. Specification of more than one simultaneous metre or key, corresponding to metrical or tonal ambiguities is therefore corresponding to music perception and also practically useful, e.g. for beat tracking systems using multiple hypotheses (e.g. Dixon, 2001).

3 Time, Metre, and Tempo

There are two representations of time in MUSITECH: metrical time and physical time. Metrical time is represented as a multiples and fractions of a pulse, where the fractions usually have a power of two as denominator. These fractions are not only used to describe durations, but also for metrical positions. Physical time is represented in microseconds since

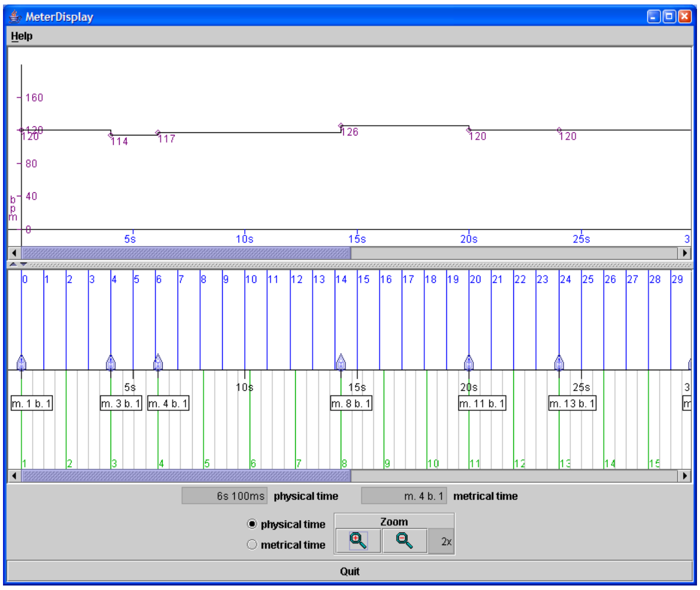

Figure 1: The MetricalTimeLine display.

the beginning of the piece (giving some reserve in accuracy compared to current playback devices). This system makes the note timing and metrical structure completely independent by making the metrical structure a view on the notes, rather than a property of them.

Having the metrical and physical time, tempo is defined by the mapping between the two times. This mapping, the *time flow* as it is named by MAZZOLA (1990), maps from the the metrical time (which Mazzola calls the musical time) to the physical time. The tempo can then be defined as the derivative of the time flow. In MUSITECH, unlike in Mazzola's approach, this mapping does not determine the timing of any actual notes, as these have their individual physical times. Notes may differ from the regular metrical position defined by the tempo, thus reflecting the view of tempo being a internal pulse established by the listener according to the structure of events heard. This pulse may adapt over time to perceived event timings, but it is not directly defined by them.

Mappings between physical and metrical time are represented by objects of a class MetricalTimeLine, which stores synchronisation points, which can be obtained by recording a human tapping or by using a beat-tracking algorithm. Between these points, the mappings are interpolated. This representation reflects the fact, that although the time flow may be defined to be continuous mathematically, we actually have only access to discrete points of synchronisation and the concept of tempo curves being arbitrary continuous functions is therefore problematic (see DESAIN and HONING, 1991a). The MetricalTimeLine contains also TimeSignatureMarkers, which define the metrical structure. There is a display available in MUSITECH for the MetricalTimeLine that displays the mapping of physical and metrical time as well as the resulting tempo (see. figure 1).

4 Score and Performance

Notes are represented by different classes of objects: ScoreNote and PerformanceNote, and the class Note, which aggregates the first two. The representation using different objects is necessary for two reasons:

1. the actual position of a note may differ from the metrical position defined by the current metrical structure
2. the mapping between performed notes and written notes is not bijective:
   - embellished notes are one note in notation but many notes in performance
   - tied notes are written as multiple notes but played as one
   - notes at one metrical time are usually played at multiple performance times, e.g. arpeggio.

The differences between the mathematical onset times of notes and the actual performance has often been referred to as 'expressive timing' (see e.g. Desain and Honing, 1991b), and this has often been interpreted as fluctuations in tempo. Yet, this is not an adequate interpretation, as notes may deviate from an expected metrical raster, without immediately changing it. This becomes also evident when several instruments play together and are not synchronous in timing without invoking an impression of two metres. Other examples are arpeggiated chords, unevenly spaced patterns like in swing style, or the 'late' timing employed by jazz soloists (see Ashley, 2002), all of which do not involve tempo changes but a variance between the metrical and the performed position of the note.

The PerformanceNote and ScoreNote contain different types of information, the ScoreNote contains diatonic pitch information, as well as onset and duration in metrical time. Therefore, tied notes are represented as several notes on the score level. PerformanceNote contains pitch information as MIDI note numbers and physical onset and duration times, but with duration information instead of MIDI's note-on and note-off events. The combination in Note objects allows 1–n and n–1 mappings to represent ornaments and tied notes. Separate score and performance objects facilitate modelling the relation between score and performance explicitly, e.g. when representing several performances of one score. It is also useful from a technical point of view, because it provides the relevant information for playback and display components in appropriate forms. Western notation uses the score notes, while piano roll display and MIDI playback use the performance notes as shown in figure 2. Tasks like the alignment of performed notes to a score have here a natural representation as an alignment of ScoreNotes and PerformanceNotes.

Additional information in the score are lyrics and harmony symbols. In music notation software these are normally treated as attributes of notes. In MUSITECH they are objects that can be positioned in a piece using a time position, either physical or metrical. Thus lyrics or harmony symbols and notes can be modelled independently, but is is also possible to establish links between them, e.g. to allow software to move them together.

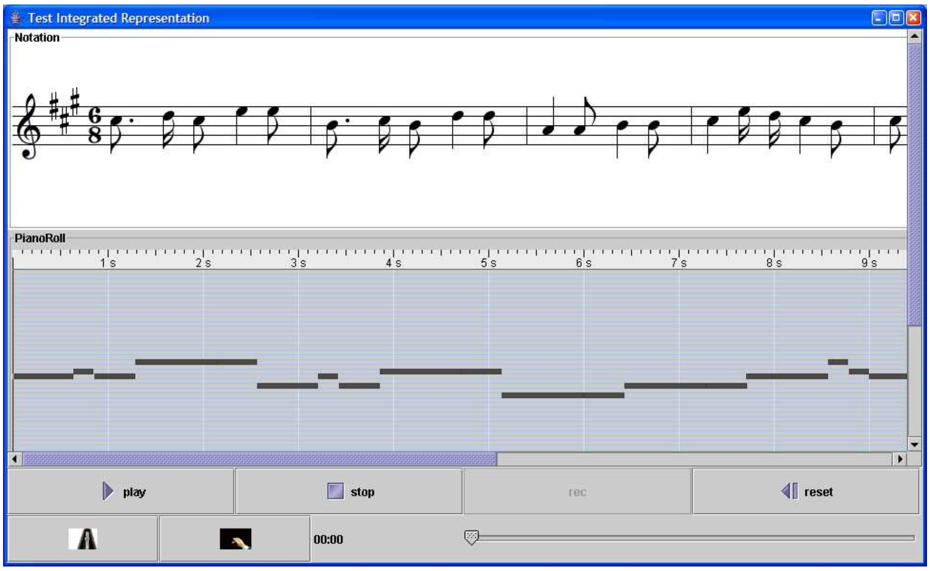

Figure 2: Display and playback of score and performance notes.

5 Structure Representation

While in music theory it is well established that in a piece of music there are several hierarchical and non-hierarchical structures in parallel (see e.g. Lerdahl and Jackendoff, 1983), most common formats cannot represent this. E.g. MIDI allows only a horizontal division, and in MusicXML (see Good, 2001) only the structure of bars and voices is possible.

The MUSITECH object model supports the specification of arbitrary sets of objects, and can therefore be used for modelling any musical structures not restricted to any specific musical or psychological theory. Thus MUSITECH supports applications in music analysis (e.g. form models) and music-psychological models (e.g. melodic segmentation). Using these structures, it is possible to define multiple parallel hierarchical (e.g. formal) and non-hierarchical (e.g. similarity-based) structures, as it is necessary to model concepts like form, voices, counterpoint, harmonic, thematic, and motivic structure.

This structure representation allows the modelling of several aspects of a composition. For instance, a sonata can have a formal structure, a thematic similarity structure, a harmonic structure, a voice structure; and all of these can be interrelated as sketched in figure 3. Some parallel hierarchies are indicated like the horizontal voice structure and the vertical form. Others are non-hierarchical graphs, like the similarity or contrast of motifs. An interesting feature is that these sets can be nested (supporting even circular structures) and connected with any other type of MUSITECH objects. Normally these structural elements will at some level refer to some visual or audible object, such as performance or score representation, but they may also represent abstract structures. The possibility of having unrestricted multiple structures makes this representation interesting for musicological repositories or databases (see Noll et al., 2002). Some research on the integration of melodic segmentation and similarity has also been performed using the this model (Weyde, 2002).

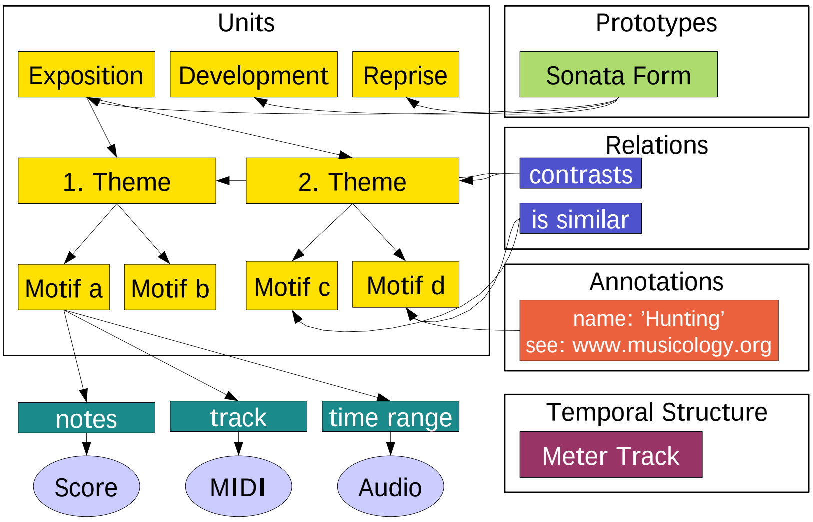

Figure 3: Different types of structures representation for a piece of music.

6 Metadata, Annotations, and Links

MUSITECH objects can be associated with metadata on all structural levels. This allows to attach categorised data, based on existing or newly defined taxonomies, as text, graphics, or hyperlinks. There is a profile mechanism, which allows MUSITECH to adopt Metadata standards like Dublin Core (see Dempsey and Weibel, 1996), and could even be used with complex metadata systems like MPEG-7 (see Martínez, 2004). These metatdata support not only information retrieval, but annotations that are important for music production, analysis, and educational applications. MUSITECH comprehensively supports annotations by the association of any object with metadata that can be used for instance for Schenkerian analysis or marking up leitmotifs.

7 Environment and Applications

In the MUSITECH project, a software infrastructure has been created, that supports the development of applications (Gieseking and Weyde, 2002). The MUSITECH representation can be stored in an XML format called MusiteXML that can also be used for transfer between applications or systems. The MUSITECH infrastructure provides a component framework and a significant number of components for display, editing, and playback have been developed, covering audio waveforms and spectra, piano roll, western notation, list-editors, as well as interactive and dynamic network-visualisations of musical structure.

Visualisation and Playback

The MUSITECH framework provides interfaces for playback and visualisation modules, and several modules have already been realised. This enables users to visualise their data and programmers to develop new visualisation or playback modules.

Services

The MUSITECH framework provides a number of services for applications. There are services for registering editing display components, playback timing, and data change, selection, and persistence. These services are transparent to the application. There are implementations that function locally; an implementation that work in peer-to-peer network are currently under development.

File Format

The object representation can be stored in an XML format called MusiteXML, that can be used for storage and transfer. In the course of the MUSITECH project we participated in the development of the Symbolic Music Representation that is currently being standardised as part of MPEG-4 (MPEG SMR)(see BELLINI ET AL., 2004). Although MUSITECH was not chosen as the Reference Model for the MPEG SMR, the requirements for the MPEG SMR standard have been influenced by MUSITECH and it is intended to make the MUSITECH software compatible to the MPEG-SMR, once it has been established.

8 Conclusions

Many musical software applications could benefit from music representation that offers more powerful ways to represent musical knowledge, as the complex analytical and cognitive structures of music are not supported by current music representations and software. The MUSITECH object model aims to support research and application development with a music representation suitable for modelling these structures. The possibilities of using and connecting multiple representation levels, flexible metre and key models, the representation of arbitrary structures, and extensive annotation capabilities support the development of advanced music applications. An implementation of this model and numerous software tools that facilitate the development of advanced musical software applications are provided in the MUSITECH infrastructure.

References

ASHLEY, RICHARD: *Do[n't] Change a Hair for Me: The Art of Jazz Rubato. Music Perception*, 19(3):311–332, 2002.

BELLINI, PIERFRANCO; BARTHELEMY, JEROME; NESI, PAOLO; and ZOIA, GIORGIO: *A proposal for the integration of symbolic music notation into multimedia frameworks.* In *Proceedings of the Fourth International Conference on Web Delivering of Music*, pp. 36–43. IEEE, 2004.

DEMPSEY, LORCAN and WEIBEL, STUART L.: *The Warwick Metadata Workshop: A Framework for the Deployment of Resource Description. D-Lib Magazine*, 2(7), 1996.
URL `http://www.d-lib.org`

DESAIN, PETER and HONING, HENKJAN: *Tempo Curves Considered Harmful. Array: Journal of the ICMA*, 11(3), 1991a.

Desain, Peter and Honing, Henkjan: *Towards a Calculus for Expressive Timing in Music Performance. Computers in Music Research*, 3:43–120, 1991b.

Dixon, Simon: *Automatic Extraction of Tempo and Beat from Expressive Performances. Journal of New Music Research*, 30(1):39–58, 2001.

Gieseking, Martin and Weyde, Tillman: *Concepts of the MUSITECH Infrastructure for Internet-Based Interactive Musical Applications.* In Busch, Christoph; Arnold, Michael; Nesi, Paolo; and Schmucker, Martin (eds.), *Proceedings of the Second International Conference on WEB Delivering of Music (WEDELMUSIC 2002)*, pp. 30–37. IEEE / Fraunhofer IGD, Darmstadt, 2002.

Good, Michael: *MusicXML: An Internet-Friendly Format for Sheet Music.* In *XML 2001 Conference Proceedings web site*. 2001. See http://www.musicxml.org.

Lerdahl, Fred and Jackendoff, Ray: *A Generative Theory of Tonal Music.* The MIT Press, Cambridge, Mass., 1983.

Martínez, José M.: *MPEG-7 Overview (version 10).* Tech. Rep. N6828, ISO/IEC JTC1/SC29/WG11, Palma de Mallorca, 2004.
URL `http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm`

Mazzola, Guerino: *Geometrie der Töne.* Birkhäuser, Basel, 1990.

Noll, Thomas; Garbers, Jörg; Höthker, Karin; ChristianSpevak; and Tillman-Weyde: *Opuscope – Towards a Corpus-Based Music Repository.* In Fingerhut, Michel (ed.), *Proceedings of the 3rd International Symposium on Music Information Retrieval 2002*, pp. 291–292. IRCAM, Paris, 2002.

Weyde, Tillman: *Integrating Segmentation and Similarity in Melodic Analysis.* In Stevens, K.; Burnham, D.; McPherson, G.; Schubert, E.; and Renwick, J. (eds.), *Proceedings of the International Conference on Music Perception and Cognition 2002*, pp. 240–243. University of New South Wales, Sydney, Australia, 2002.