Social Media Retrieval for Music Education

Mathieu Barthet, György Fazekas, Simon Dixon, Mark Sandler
Centre for Digital Music, Queen Mary University of London
Mile End Road, London E1 4NS, UK
{mathieu.barthet, gyorgy.fazekas, simon.dixon, mark.sandler}@eecs.qmul.ac.uk

ABSTRACT

Advances in content-based multimedia analysis, recommender systems and Web-based social platforms for content and metadata sharing provide opportunities to create novel applications for music education. In this paper we describe a framework for intelligent music tutoring systems, through the combined use of content and context-based approaches. First, we investigate traditional computer-assisted music education applications, and review music information retrieval and Web technologies relevant to social media retrieval and music education. We discuss semantic aspects of these technologies and the use of ontologies as common grounds for structuring heterogeneous information available on the Web and from machine analyses. The importance of multimodality in music education tools is highlighted before we discuss how the reviewed technologies and information resources may be combined in interactive tools for music learning, for instance, a tool for searching the Web for guitar tablatures and YouTube video tutorials.

Categories and Subject Descriptors

H.5.5 [**Sound and Music Computing**]: [Signal analysis, synthesis, and processing, Modeling]; H.3.3 [**Information Search and Retrieval**]: [Retrieval models]; J.5 [**Computer Applications**]: Arts and Humanities—*Music*

1. INTRODUCTION

Although there have been numerous applications of computers for music education [3], the recent explosion of digital media collections, resources, analysis and classification technologies offer the potential to revolutionise the data retrieval and instructional strategies on which they are drawn. Online education (so-called e-learning) is a burgeoning area in remote education with high economic implications and promises to attract and benefit a wide audience. This paper aims to show how social media retrieval (SMR) can foster the development of intelligent music tutoring systems (IMTS) for music learners, through the combined use of content and context-based media analysis. The application of digital music technology in music education is an important topic of theory research in the field of music teaching [14]. Indeed, with the rapid growth and development of information and communication technologies (ICT), there is a need to investigate whether ICT can be efficiently integrated into instruction and learning processes. Results of a recent study [12] showed that IT-assisted music teaching can improve music teaching quality. With the popularity of Web platforms such as YouTube revolutionising the way to consume, create and share music tracks and video clips [4], as well as online streaming services like Last.fm, it is expected that e-learning will incorporate more multimedia-based learning tools targeted for the average user, rather than learning tools for traditional educational institutions [8]. Figure 1 presents the proposed framework for an e-learning music education service relying on social media retrieval and the "triple synergy paradigm" (content/metadata, user/user and user/content interactions) [10]. IMTS relying on social media retrieval could hence benefit from social relations and interactions between music learners. For instance, students having similar musical taste or similar skill-level may learn from a same set of songs. Therefore the combination of social and hybrid recommendation technologies with educational tools could bring significant advances in this emerging field.

2. CONTENT-BASED ANALYSIS

High-level semantic features, closely corresponding to musically meaningful concepts (e.g. notes, rhythm, tempi, keys, chords or structural segments) may be extracted either automatically or semi-automatically relying on content analysis, potentially fused with crowd-sourced metadata available on the Web. Content-based audio analysis aims at extracting information from signals, a requisite for creating interactive educational tools for music learning [1], [9]. Ethnographic observations of musicologists have shown that content-based analysis and visualisations helped the understanding of subtle performance practices, which could not be reached by only relying on closed listening [2]. Figure 2 shows how social data, content-based analysis and expert knowledge can be combined to recommend music adapted to user skills. To build adaptive song recommendation systems for music learning, digital music sheets and tabs are also interesting resources since information to classify songs according to musical concepts that are useful in a learning curriculum context (e.g. harmony, melodic line, rhythmic patterns) can potentially be inferred from them.

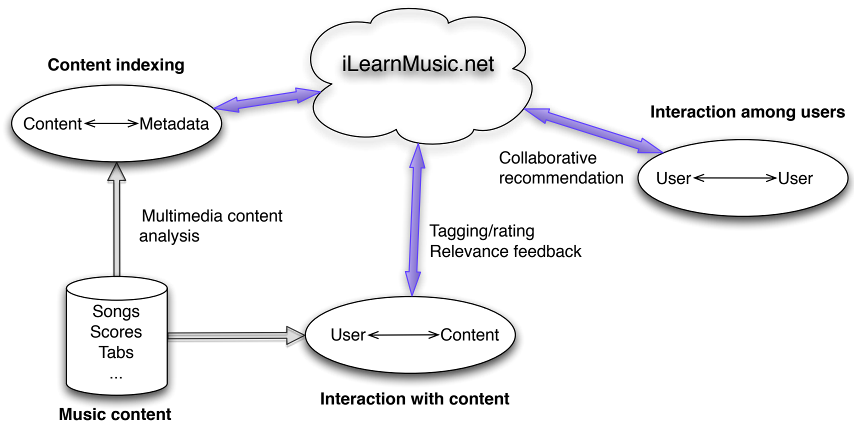

Figure 1: Example of e-learning music education service based on the "triple synergy paradigm" [10].

3. ASPECTS OF SEMANTICS

Prevailing machine learning tools provide good solutions to particular problems, however they do not easily allow us to close the semantic gap between features and computational models on one hand, and musicological descriptors or human music perception on the other. While cognitive modelling, the use of contextual metadata, and the use of high-level reasoning are promising future directions [19], some common agreement in how knowledge and information are represented is requisite for deploying complex systems [5]. The *linked open data* movement aims at making data freely available on the Web, which can be used in music recommendation and learning. The published data sources cover a wide range of topics: from music (Musicbrainz, Magnatune and Jamendo) to encyclopedic information or bibliographic information (Wikibooks, DBLP bibliography). These datasets are published in a way that follows common data modelling principles and technologies. Semantic Web ontologies provide the foundations for combining heterogeneous data sources in the linked data cloud, as well as a promising way of representing knowledge and information in a uniform and interoperable way in complex systems. The Music Ontology [15] serves as a highly adopted base ontology for the music domain. It provides fundamental concepts and relationships within well-defined boundaries, while these fundamental concepts are used as basis for richer knowledge representations provided in domain specific extensions. The core of the Music Ontology can be used to describe the music production workflow (e.g. the process of composition, performance and recording), and allows to make social links at each level. The second important aspect is its extensibility: several sub-ontologies were produced for describing music similarity [11], audio analysis algorithms and their output features [7], musicological data such as chord progressions, keys or musical temperament, musical instruments [13] and an ontology of music production in the studio [6]. An ontology that allows collaborative annotation of music performances, grounded on the Music Ontology, is presented in [17].

4. MULTIMODALITY

Multimodality has shown to facilitate the negotiation of musical concepts between teachers and students in the context of music lessons [18], and to provide an increased access to meaning by combining several types of feedback while listening (aural, visual, textual, etc.) in the context of musicological research [2]. In [16], combining audio and video feedback and hand motion analysis is proposed to provide piano performers with extensive information about the physiology of music performance. The authors demonstrate that such multimodal feedback can be helpful in identifying and correcting technical problems and resolving repetitive stress injuries in pianists. We suggest that intelligent tutoring systems relying on social media retrieval should be multimodal with several types of feedback combining e.g. audio, video, and textual/symbolic streams.

5. CASE STUDY

Hotttabs, a multimedia guitar tutor detailed in [1] integrates several of the functionalities described throughout this article in a client-server architecture. The client provides a Web interface to perform the following actions: *i)* retrieve popular song recommendations, *ii)* retrieve a list of YouTube model performance videos, and three sets of guitar tabs clustered by complexity (easy, medium, and difficult) for a selected song, *iii)* view video tutorials synchronised with automatically extracted chords, *iv)* follow a link to a tab, as one would do using a search engine. Hotttabs relies on a multimodal approach. It uses crowd-sourced tablatures mined from the Web using its built in search engine and tab format parser, popular song recommendations provided by The Echo Nest, video tutorials accessed via the YouTube application programming interface (API), and chord visualisations extracted automatically from user generated content. For instance, it is possible to choose guitar tabs and model videos that are most appropriate for the users' skill level. This choice is facilitated by the automatic chord extraction process. Song selection is made easier using popular song recommendation, while the choice of songs is not limited to a predefined catalogue, thanks to the large number of tabs and tutorials created and published by the community. Hotttabs integrates several data sources by aggregating and internally correlating information in a process whereby the Music Ontology provides the common ground. However, due to the lack of machine-processable information in many

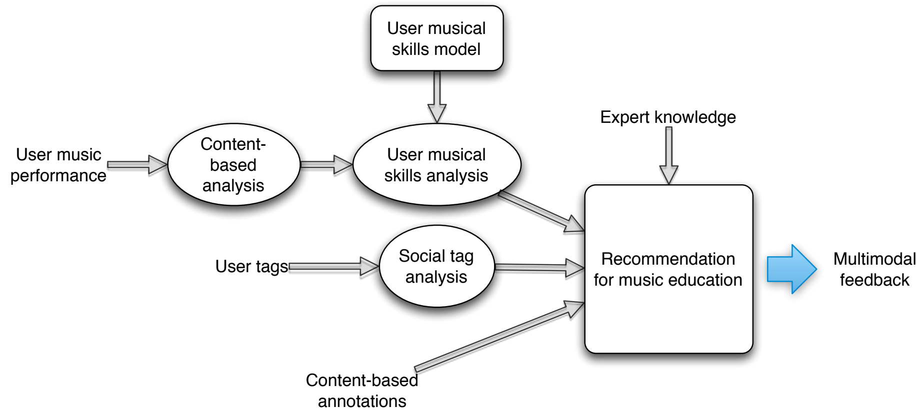

Figure 2: Adaptive music learning framework based on social media retrieval and user skills modeling.

social sites (of which guitar tab sharing sites are prime examples) and the disharmonious nature of Web APIs as well as their underlying data models that are commonly defined in an ad-hoc manner, building educational applications using social media retrieval still presents a very significant challenge. This process could be made easier, if social media sites were following the principles advocated by the Linked Data community, and proprietary Web APIs were replaced by dynamic query interfaces grounded in shared ontology-based schemata.

6. CONCLUSIONS

In this paper we introduced a new framework for intelligent music education systems relying on social media retrieval and music recommendation, and presented a first application for guitar learning. Further research is needed to find out how to model user skills to make meaningful recommendation for learning. Automated ways of providing cognitive feedback based on the results of performance practice analysis using content-based methods is an additional promising topic of research. Although successful computer applications can already be demonstrated using current technologies, we acknowledge the importance of the social aspects of music-making and agree that computer-assisted music education needs to be complemented by face-to-face interactions with other musicians and/or teachers.

7. ACKNOWLEDGEMENTS

This study was conducted as part of the RCUK Digital Economy project EP/I001832/1.

8. REFERENCES

[1] M. Barthet, A. Anglade, G. Fazekas, S. Kolozali, and R. Macrae. Music recommendation for music learning: Hotttabs, a multimedia guitar tutor. In *Proc. WOMRAD*, 2011.

[2] M. Barthet and S. Dixon. Ethnographic observations of musicologists at the british library: Implications for music information retrieval. In *Proc. ISMIR*, 2011.

[3] M. Brandao, G. Wiggins, and H. Pain. Computers in music education. In *Proc. AISB'99 Symposium on Musical Creativity*, 1999.

[4] C. Cayari. The youtube effect: How youtube has provided new ways to consume, create, and share music. *Int. J. of Education and the Arts*, 12(6):1 – 30, 2011.

[5] G. Fazekas, Y. Raimond, K. Jakobson, and M. Sandler. An overview of Semantic Web activities in the OMRAS2 Project. *J. of New Music Research*, Vol. 39 issue 4:pp. 295–311, 2010.

[6] G. Fazekas and M. Sandler. The Studio Ontology Framework. In *Proc. ISMIR*, 2011.

[7] G. Fazekas, T. Wilmering, and M. Sandler. A knowledge representation framework for context-dependent audio processing. In *Proc. 42th Int. Conf. of the AES on Semantic Audio*, 2011.

[8] G. Friedland, W. Hurst, and L. Knipping. Educational multimedia. *IEEE Multimedia*, 15(3):54 – 56, 2008.

[9] S. Grollmisch, E. Cano, and C. Dittmar. Songs2see: Learn to play by playing. In *Proc. AES Int. Conf. on Audio for Games*, 2011.

[10] A. Hanjalic. Social media retrieval. In *Summer School on Multimedia Semantics*, Amsterdam, 2010.

[11] K. Jacobson, Y. Raimond, and M. Sandler. An ecosystem for transparent music similarity in an open world. In *Proc. ISMIR*, 2009.

[12] Y. Ji and T. Chen. Improve music teaching quality throughout using information technology. In *Proc. ISDPE*, pages 90 – 94, 2010.

[13] S. Kolozali, G. Fazekas, M. Barthet, and M. Sandler. Knowledge Representation Issues in Musical Instrument Ontology Design. In *Proc. ISMIR*, 2011.

[14] R. Qionggang. Application and research on digital music technology in music teaching. In *Proc. IEEE Int. Conf. on Computer-Aided Industrial Design Conceptual Design (CAID CD)*, pages 1947 – 1949, 2009.

[15] Y. Raimond, A. Samer, M. Sandler, and F. Giasson. The music ontology. In *Proc. ISMIR*, 2007.

[16] K. Riley, E. E. Coons, and D. Marcarian. The use of multimodal feedback in retraining of complex technical skills of piano performance. *Medical Problems of Performing Artists*, 20(2):82 – 88, 2005.

[17] V. Sebastien, D. Sebastien, and N. Conruyt. An ontology for musical performances analysis. In *Proc. Int. Conf. on Internet and Web Applications Services*, 2010.

[18] D. Stowell and S. Dixon. MIR in School? Lessons from Ethnographic Observation of Secondary School Music Classes. In *Proc. ISMIR*, 2011.

[19] J. Wang, X. Chen, Y. Hu, and T. Feng. Predicting high-level music semantics using social tags via ontology-based reasoning. In *Proc. ISMIR*, 2010.