jMIR and ACE XML: Tools for Performing and Sharing Research in Automatic Music Classification

Cory McKay
Schulich School of Music
McGill University
Montreal, Quebec, Canada
514-398-4535 x0300
cory.mckay@mail.mcgill.ca

John Ashley Burgoyne
Schulich School of Music
McGill University
Montreal, Quebec, Canada
514-398-4535 x0300
ashley@music.mcgill.ca

Ichiro Fujinaga
Schulich School of Music
McGill University
Montreal, Quebec, Canada
514-398-4535 x00944
ich@music.mcgill.ca

ABSTRACT

This presentation will begin by introducing the research fields of music information retrieval and automatic music classification. The core of the presentation will then be divided into two parts, the first dealing with the jMIR software suite, and the second dealing with the ACE XML file formats. jMIR is a set of free and open source software tools for automatically classifying music in a variety of ways. ACE XML is set of XML-based standardized file formats for storing and sharing the essential information that is related to automatic music classification. Although ACE XML is supported by each of the jMIR components, it is intended for use as a general standard in automatic music classification, and is not limited specifically to jMIR.

Categories and Subject Descriptors

E.5: [**Files**]: Organization/structure

I.5.4 [**Pattern Recognition**]: Signal processing

J.5 [**Arts and Humanities**]: Music

General Terms

Algorithms, experimentation, standardization, languages.

Keywords

Music information retrieval, automatic music classification, data representation, machine learning, pattern recognition, feature extraction, audio, symbolic music, web mining.

1. INTRODUCTION

Music information retrieval (MIR) is a field of research that concentrates on computationally extracting useful information from music and making this information accessible to users of all kinds in ways that meet each of their particular needs. MIR overlaps with and is enriched by many other disciplines, including machine learning, data mining, digital signal processing, music theory, musicology, music psychology and the library sciences.

There are three primary types of musical information that are traditionally considered in MIR: audio musical recordings (e.g., MP3s, WAVs and FLACs); symbolic musical representations (e.g., scans of scores, MIDI files, OSC files and Humdrum files); and cultural information that is available on the Internet and elsewhere (e.g., edited metadata repositories, unedited listener tags, sales statistics, playlists and web co-occurrences). Other sources of information, such as lyrics and album art, can also be analyzed by MIR researchers as well.

Many area of MIR research utilize machine learning techniques to classify music in some way. This classification can sometimes involve relatively low-level classifications, such as differentiating note onsets from sustained notes or silence in an audio recording. Such low-level classifications can then in turn be used iteratively to perform increasingly high-level classifications, such as classifying the pitch of notes in audio waves, using these pitches to classify chords and then using these chords to classify structural harmonies. Other types of classification can begin with high-level classes, such as classification by genre, mood or geographical origin; or identification of performer, composer or song title. Still other types of classification may be tailored to individual users, such as automatic music recommendation and playlist generation, or to large demographic groups, such as hit prediction and user tag prediction.

Most MIR classification research involves three core tasks:

- Collecting and annotating ground truth data for training and testing classification algorithms.
- Extracting features (i.e., characteristic pieces of information) from the data.
- Using pattern recognition algorithms to associate feature patterns with particular classes.

Ideally, one would like to have standardized, general, flexible, extensible and easy-to-use software tools and file formats designed for MIR research in general, and these tasks in particular. This would not only facilitate the process of comparing, evaluating, sharing and building upon the approaches used by other researchers, but would also significantly reduce duplicated effort and the associated slowdown in the progress of MIR research.

JCDL '09, June 15–19, 2009, Austin, Texas, USA.

Unfortunately, such MIR software and file formats are currently lacking. Although there are certainly a number of excellent MIR software systems (e.g., CLAM [1], MIDIToolbox [2], MIRToolbox [3] and Marsyas[12]), none of them address all three of the core automatic classification tasks described above with respect to audio, symbolic and cultural musical data. With respect to file formats, the commonly used Weka ARFF [13] format has some significant limitations with respect to the needs of music classification [5]. Although the alternative Music Ontology [11] representation is very powerful with respect to the representation of general musical ontologies, it is not designed for the specific needs of automatic classification.

The jMIR software and the ACE XML file formats are presented as solutions to these needs. They hold potential not only as tools for facilitating the performance and sharing of MIR research, but also as resources that are available for free to users of all kinds, including but certainly not limited to music librarians.

2. jMIR

jMIR is a suite of standardized software for use in automatic music classification research:

- *jAudio [4]:* An audio feature extractor that includes implementations of 26 core features. jAudio also includes implementations of “metafeatures” and “aggregators” that can be used to automatically generate many more features from these core features (e.g., standard deviation, derivative, etc.).
- *jSymbolic [6]:* A symbolic feature extractor for processing MIDI files. jSymbolic is packaged with 111 implemented features.
- *jWebMiner [8]:* A cultural feature extractor that extracts features from the web based on search engine co-occurrence page counts. Many user options are available to improve results, including search synonyms, filter strings and site weightings.
- *ACE [5]:* A meta-learning classification system that automatically experiments with a variety of different dimensionality reduction and machine learning algorithms in order to evaluate which are well suited to particular problems. ACE can also be used as a simple automatic classification system.
- *Codaich, Bodhidharma MIDI* and *SAC [7]:* Collections of, respectively, MP3 files; MIDI files; and matched MP3, MIDI and cultural data. The recordings are all labeled with ground truth metadata.
- *jMusicMetaManager [7]:* Software for managing large audio data collections. Extracts metadata from iTunes XML files or MP3 ID3 tags and then produces 39 different reports that statistically profile the dataset and detect metadata errors, inconsistencies and redundancies.

The jMIR components can be used either independently or as an integrated suite, and are intended for users of varying technical backgrounds. Most of the components include a GUI and electronic manual, and all of the components are entirely open source.

jMIR is intended not only for direct use in automatic music classification, but also as a framework for collaboratively developing new MIR approaches. Each jMIR component is based upon a plug-in architecture and is designed to be highly extensible. For example, one researcher might design a new feature that can be extracted from audio files. This feature could be implemented as a plug-in to jAudio, and could then be distributed to other researchers who use jAudio, who could in turn use this feature to iteratively develop new features based on this original feature.

jMIR is also designed to facilitate and promote research combining audio, symbolic and cultural data. The approach of using more than one such data type at a time has been found experimentally to be advantageous [9].

3. ACE XML

ACE XML [10] is a set of standardized file formats designed for storing and sharing data associated specifically with automatic music classification. This includes the representation of feature values extracted from instances; class labels and metadata associated with instances and their sub-sections; abstract feature descriptions and parameterizations; and class ontologies and metadata.

These formats have been designed to address the significant shortcomings in the file formats currently most commonly used in MIR research. To provide just a few examples, ACE XML makes it possible to associate multiple class labels with a single instance, to specify relationships between class labels, to group associated feature values in ways that can be meaningful to machine learning algorithms and to maintain associations between instances and their sub-sections and metadata, none of which can be done conveniently with the file formats that are currently typically used, such as Weka ARFF.

A new update to ACE XML, ACE XML 2.0, introduces many improvements to the original ACE XML 1.1 formats. Examples include the abilities to specify links to external resources using RDF-like triples, to specify weighted class memberships, to express feature arrays of arbitrary dimensionality and to reduce file sizes using compression.

All of the jMIR components can read and write ACE XML, although they can also be used with Weka ARFF files if necessary. ACE XML 2.0 is designed for use well beyond the specific scope of jMIR, however, and is proposed for general adoption in both automatic music classification research and practical use.

4. ACKNOWLEDGMENTS

The authors would like to thank the Andrew W. Mellon Foundation for their generous financial support as well as the NEMA group members for their valuable ideas and suggestions.

5. REFERENCES

[1] Arumi, P., and X. Amatriain. 2005. CLAM, an object oriented framework for audio and music. *Proceedings of the International Linux Audio Conference.*

[2] Eerola, T., and P. Toiviainen. 2004. MIR in Matlab: The MIDI Toolbox. *Proceedings of International Conference on Music Information Retrieval*. 22–7.

[3] Lartillot, O., P. Toiviainen, and T. Eerola. 2008. A Matlab toolbox for music information retrieval. In *Data Aalysis, Machine Learning and Applications,* ed. C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker. New York: Springer. 261–8.

[4] McEnnis, D., C. McKay, and I. Fujinaga. 2006. jAudio: Additions and improvements. *Proceedings of the International Conference on Music Information Retrieval*. 385–6.

[5] McKay, C., R. Fiebrink, D. McEnnis, B. Li, and I. Fujinaga. 2005. ACE: A framework for optimizing music classification. *Proceedings of the International Conference on Music Information Retrieval*. 42–9.

[6] McKay, C., and I. Fujinaga. 2006. jSymbolic: A feature extractor for MIDI files. *Proceedings of the International Computer Music Conference.* 302–5.

[7] McKay, C., D. McEnnis and I. Fujinaga. 2006. A large publicly accessible prototype audio database for music research. *Proceedings of the International Conference on Music Information Retrieval*. 160–3.

[8] McKay, C., and I. Fujinaga. 2007. jWebMiner: A web-based feature extractor. *Proceedings of the International Conference on Music Information Retrieval*. 113–4.

[9] McKay, C., and I. Fujinaga. 2008. Combining features extracted from audio, symbolic and cultural sources. *Proceedings of the International Conference on Music Information Retrieval.* 597–602.

[10] McKay, C., J. A. Burgoyne, J. Thompson, and I. Fujinaga. 2009. Using ACE XML 2.0 to store and share feature, instance and class data for musical classification. Submitted to the *International Society for Music Information Retrieval Conference*.

[11] Raimond, Y., S. Adbdallah, M. Sandler, and F. Giasson. 2007. The Music Ontology. *Proceedings of the International Conference on Music Information Retrieval.* 417–22.

[12] Tzanetakis, G., and P. Cook. 2000. Marsyas: A framework for audio analysis. *Organized Sound* 10: 293–302.

[13] Witten, I., and E. Frank. 2005. *Data mining: Practical machine learning tools and techniques with Java implementations.* San Francisco: Morgan Kaufmann.