Supporting the Music Information Retrieval Research Community – a Use Case for the Semantic Grid

David De Roure

Electronics and Computer Science
University of Southampton, UK

dder@ecs.soton.ac.uk

J. Stephen Downie

Graduate School of Library
and Information Science
University of Illinois at Urbana-Champaign

The Music Information Retrieval research community is a multidisciplinary community working together on techniques for the extraction of features from musical content. These techniques can involve significant computation, and there is a requirement for remote execution due to content access issues. Additionally, there is a significant need to generate and share annotations and derived data within the community. The techniques of Semantic Grid appear well suited to delivering the computational, data and collaborative infrastructure required by this community, and this application is innovative with respect to other Semantic Grid projects in adopting a 'social tagging' approach. Additionally, this study suggests a novel application of Semantic Grid in software component description and the evaluation of software and algorithms.

Introduction

Feature extraction takes the information contained in multimedia data and creates an abstraction of its content, often called a signature or feature vector. For example, in the case of images the features could be colour and texture. Feature extraction techniques provide a mechanism for analysing multimedia content, and they also support querying of content by working with the abstraction rather than the content itself, leading to their use in search engines and content-based retrieval systems.

Music Information Retrieval is then the science of extracting features from musical content, such as melody and tempo, to facilitate tasks such as analysis and music retrieval. This research community is highly interdisciplinary, ranging from signal processing experts to music scholars and digital libraries experts. [1]

The community has established an array of software tools to support this work (see http://www.music-ir.org).

MIREX

In 2005 the 1st Music Information Retrieval Evaluation eXchange (MIREX) took place during the 6th International Conference on Music Information Retrieval (ISMIR). MIREX aims to compare state-of-the-art algorithms and systems relevant for Music Information Retrieval. The evaluations focused on the following:

Audio Artist Identification
Audio Drum Detection
Audio Genre Classification
Audio Melody Extraction
Audio Onset Detection
Audio Tempo Extraction
Audio and Symbolic Key Finding
Symbolic Genre Classification
Symbolic Melodic Similarity

To achieve this, people submit their code to a central site where it is executed against a standard database of musical content. The centralisation is necessary because the content itself cannot be distributed, for licensing reasons.

Requirements

Running MIREX has raised a number of technical challenges which this community wishes to address, including:

- Ease of execution of remote code by the community (a Web Services solution is being considered);
- The computational challenge (the codes can take a long time to run);
- Sharing of intermediate data used by the algorithms (since some algorithms perform the same pre-processing stages – for example, Fast Fourier Transforms);
- Creation and sharing of annotations: establishing the ground truth against which the algorithms are compared (currently achieved using experts but this could extend to the broader community) ; and sharing of automatically extracted annotations; sharing of performance results from running software components and algorithms.

The last point is significant, since the annotations are formed by a diverse community of users. This is an example of 'social tagging', where the community create the metadata, and is an area of metacontent creation which is becoming increasingly significant. The collaborations within the community, and the sharing of resources and services, can be seen as formation of social networks or virtual organisations.

For people who want to make use of the algorithms in their research, we envisage a future research environment in which there may be multiple services available which perform a variety of feature extraction functions. Additionally, we envisage many applications of these techniques in professional and consumer systems, which raise a related set of requirements outside of the research context. Application scenarios include:

- Content based retrieval – taking a clip of music and finding pieces that match;
- Content-based navigation – automatic generation of hypermedia links in musical content by defining the anchors as features which are matched against the content [2];
- Finding suitable pieces of music based on specification of certain features (such as genre and tempo), for example in multimedia authoring;
- Analysing, organising and generating personal playlists for portable audio players;
- Music recommender systems based on musical selections (e.g. collections, playlists) of multiple users;
- Detection of plagiarism of musical content.

Towards Semantic Grid solutions

These challenges exhibit some classic Grid requirements – accelerated computation, working with large volumes of data, and remote execution. However the data sharing and annotation are classic Semantic Grid problems, as is the description and discovery of resources and services within the research or commercial environment. The MIREX exercise involves evaluation of algorithms and generation of intermediate results which provide additional forms of information within the system.

There are interesting distributed and collaborative dimensions, including a social tagging perspective which is previously unexplored within Semantic Grid projects. If multiple people create annotations about the same pieces of music then it should be possible for this knowledge to accumulate, providing a rich network of metacontent which is itself the basis

for new functionality (e.g. in search). To achieve this requires establishing ways of uniquely identifying content and sharing those identifiers, and describing the relationships between the "same piece of music" appearing in various forms.

At the level of describing pieces of music there are already classifications in use; for example, in online music catalogues. These vary – there is not a unique underlying conceptualisation of this domain.

The steps needed to apply Semantic Grid in this context would involve establishing a shared identifier scheme for musical content and descriptions of the relationships between musical content, and establishing several metadata schema:

1. to describe derived data so that it may more easily be found, shared and utilised;
2. to describe features (ground truth annotations and extracted features);
3. to establish metadata schema to describe feature extraction techniques and their evaluation, as in the MIREX exercise;
4. to describe software components;
5. to describe deployed services.

The third and fourth points are perhaps an area of Semantic Grid that has not been explored before – the use of Semantic Web technologies to describe software components, algorithms, tests and results, within the development phase of Grid applications.

Preliminary study

An approach that is currently being explored to take these ideas forward is the provision of an RDF (W3C Resource Description Framework) triplestore containing relevant metadata which has been harvested from available sources. Currently these the internet CD database, some personal playlists, a subset of information available from the MIREX exercise and tool information obtained from online information in the Music IR community. This diverse and changing set of information benefits from the use of triplestore approach, as opposed to using relational databases for example, due to the flexible schema. We are using the 3store RDF store, which has so far assimilated large volumes of metadata without difficulty . We anticipate that this exercise will bring to light some of the issues in realising this use case.

References

[1] Joe Futrelle and J. Stephen Downie. (2002). Interdisciplinary communities and research issues in music information retrieval. In Michael Fingerhut (Editor), Proceedings of the Third International Conference on Music Information Retrieval: ISMIR 2002. (pp. 215-221).

[2] De Roure, D. C. and Blackburn, S. G. (2000) Content-based navigation of music using melodic pitch contours. Multimedia Systems 8(3) pp. 190-200.

Acknowledgements

This exercise arose as a collaboration formed under the GGF HASS (Humanities, Arts and Social Sciences) Research Group. Thanks to Allison Clark and to the Worldwide Universities Network for facilitating this.