Large Multimedia Archives for Music-related Cultural Heritage
Adriano Baratè, Goffredo Haus, Luca A. Ludovico
Laboratorio di Informatica Musicale (LIM)
Dipartimento di Informatica e Comunicazione
Università degli Studi di Milano
Via Comelico 39/41 – I-20135 Milano, Italy
{barate,haus,ludovico}@dico.unimi.it

Abstract. This paper illustrates a number of active projects about music and cultural heritage in Italy. First, the matter of designing and implementing a multimedia database oriented to music will be introduced. Then, the paper will address the advantages of the adoption of IEEE 1599 standard inside a multimedia database. The latter aspect represents a new perspective to interact with music contents, and it can enrich the models of music fruition currently available on the Web.

1. Introduction

Nowadays large multimedia archives are commonly available for Web users. Music in particular is a field where centralized as well as distributed repositories provide a huge amount of information, in terms of both digital objects and catalogue metadata. For instance, as regards the Italian music cultural heritage, we can cite an opera house such as Teatro alla Scala, a music publisher with an important historical archive such as Ricordi, and finally Discoteca di Stato, the Italian institutional archive of audio documents: they all provide either full or authenticated access to their archives through Web-based interfaces, and their repositories are rich in digital audio objects and music-related documents.

At Laboratorio di Informatica Musicale (LIM, Università degli Studi di Milano), both the Web interface to access the mentioned archives and the underlying multimedia databases have been developed and implemented. One of the most interesting characteristics consists in providing original navigation paths, which do not force the user to adopt a tree view of the information. Some examples will be presented in the following.

Another matter – faced both from a theoretical and from an implementative point of view – consists in widening the perspective of music-related databases by unifying such information within a central node of a network. This is the aim of InternetCulturale project, which has been updated by introducing an evolved search engine for music-related contents. The InternetCulturale portal now allows a distributed query on materials coming from heterogeneous institutions, such as Archivio Storico Ricordi and Discoteca di Stato.

Finally, this paper will illustrate the advanced possibilities coming from the encoding of music contents in a new XML-based format known as IEEE 1599. The main aspect of novelty is providing to the Web user an interaction with database contents that goes beyond the mere visualization of metadata. Often, music-oriented databases contain a great amount of heterogeneous but correlated information about a music piece. When we aim at a comprehensive description of music, different materials are involved: scores and other symbolic representations of music, audio and video recordings, fliers, playbills, posters, photos, sketches, fashion plates, costumes and related accessories, stage tools, maps and equipment, and other text documents commonly used in the evening’s programs, such as bibliography, discography, libretto, short descriptions, and reviews of music works. This list does not claim completeness, but is wide enough to illustrate the heterogeneity of data and metadata that can be related to a single work. All documents give a contribution to the overall description of the music piece, and there is considerable cultural, scientific, and commercial interest in providing access to such information, as shown by the growing amount of digitization projects and multimedia databases. Please note that a database such as the one of Teatro alla Scala contains most of the mentioned materials.
2. Design and Implementation of a Multimedia Archive

An effective and efficient multimedia database presents many advantages for opera houses, conservatories, music institutions and so on. For example, it makes management operations and editorial initiatives much simpler; however, such database can assume a very important role also from the cultural heritage perspective. In fact, thanks to computer-based technologies, database contents can be easily transmitted, both in space and in time dimension.

A digital version of the original material can be handed on from generation to generation. In a certain sense, not only data and metadata, but also physical objects can be transmitted, after ad hoc digitalization campaigns.

Besides, thanks to network technologies, a worldwide access to database contents can be granted to any authorized user. On-line publication of multimedia contents requires a well structured interface, containing effective navigation tools and appropriate forms of data aggregation. As a result of this effort, researchers, performers, music students, and untrained people could access such cultural treasure, anytime and anywhere.

After choosing a suitable database management system (DBMS), possibly supporting advanced features for multimedia management, the database itself should be designed and implemented. Its schema must take into account the heterogeneity of music and music-related contents.

Let us introduce the case study of Teatro alla Scala [1]. An opera house is a very rich and complex informative environment, whose materials and documents include:

- Scores and symbolic representations of music;
- Audio recordings;
- Video recordings;
- Fliers, playbills and posters;
- Photos;
- Sketches;
- Fashion plates;
- Costumes and related accessories;
- Stage tools and equipment;
- Stage maps;
- Other textual documents, such as bibliography, discography, libretto, short descriptions and reviews of music works.

First, the original form of the material often has to be converted, as in a multimedia database only digital objects and data can be inserted. Some contents from the list are already digital encodings: for instance, digital photos, digital recordings, and computer-edited texts. Besides, some contents undergo an analog-to-digital conversion which prevents (or should prevent, under particular conditions) informative loss; it is the case of image scanning, analog audio and analog video digitalization, and typing of digital documents (texts and scores) from a hard copy. Finally, there are physical objects that can be captured only from certain points of view and a limited number of times. In this case, we know that the digital copy does not allow to catch and appreciate all the facets of the digitalized object. Let us cite the examples of a stage tool or a costume: many digital photos can be taken and stored to show their features, but the whole photographic sequence is not sufficient to provide a three-dimensional view of the original object. For physical objects, a virtually exhaustive set of digital descriptions (both textual and visual) should be chosen; but such choice would represent – in any case – an arbitrary operation which introduces informative loss. For instance, 3 photos of a costume can show its characteristics better than a single shot, as well as 100 photos give us a far better three-dimensional view of the subject, but neither 1000 photos can make up for the sensation of touching the costume’s
fabric. About this matter, we could mention some in-use techniques to provide 3D renderings of physical objects; nevertheless, they can be rarely applied to a database environment.

Figure 1 shows a simplified version of the database structure. For further details, please refer to [1]. After structuring a database in a suitable way (depending on the characteristics of the original archive), a Web interface to access information has to be designed. In this sense, the main novelties involve the navigation paths proposed to the Web user. Music-related information, originally tree-structured, can be browsed also by using non-linear paths. For instance, let us consider the screenshots presented in Figure 2. They show an example of navigation through the digital contents of Teatro alla Scala: the user first search for an opera (e.g. “Tosca”), then access a type of related multimedia objects (e.g. “sketches”), and finally browses all the sketches realized by the considered artist (e.g. “Margherita Palli”).

The characteristics we have described do not require particular techniques or technologies, rather they simply represent a new way to access information. On the contrary, advanced features for our proposal will be introduced in the next section.
At LIM, researchers have recently carried out a project to integrate the results of digitalization campaigns of Italian music materials within a unique portal, which can be seen as the central node of a network of music-related institutions. This project implements the music search engine now available through InternetCulturale Web site, a portal managed by the Italian Ministry for Cultural Heritage and Activities (see Figure 3). The address is: http://www.internetculturale.it.
3. Music-related archives and IEEE 1599

IEEE 1599-2008 is a new file format whose international standardization was achieved on September 26, 2008. Its ultimate goal is to provide a highly integrated representation of music, where score, audio, video, and graphical contents can be appreciated together.

This standard provides an XML-based format. There are many advantages in choosing XML to describe information in general, and music information in particular, as the large number of other music-oriented XML languages demonstrates. For instance, an XML format is typically open, free, cross-platform, easy to read (both by humans and by computers), well-supported by programming languages and parsing software. Besides, XML fits well to music hierarchical structures, moreover it is flexible and extensible as required by traditional and contemporary music languages. All these topics have been treated in [2], [3] and [4].


Now let us introduce the key features of the standard. A comprehensive description of music must support heterogeneous materials. Thanks to the intrinsic capability of XML to provide strongly structures for information, such representations can be organized in an effective and efficient way. IEEE 1599 employs six different layers to represent information:

- General – music-related metadata, e.g. title, author, genre, and other catalogue information about the piece;
- Logic – the description of score in terms of music symbols;
- Structural – identification of music objects and their mutual relationships;
- Notational – graphical representations of the score;
- Performance – computer-based descriptions and executions of music according to performance languages;
- Audio – digital or digitized recordings of the piece.

Needless to say, not all layers must, or can, be present for a given music piece. Of course, the higher their number, the richer the musical description.

Richness has been mentioned in regard to the number of heterogeneous types of media description, namely symbolic, logic, audio, graphic, etc. But the approach in IEEE 1599 allows one extra step,
namely, that each layer can contain many digital instances.

For example, the Audio layer could link to several audio tracks, and the Structural layer could provide many different analyses for the same piece. The concept of multi-layered description (i.e. as many different types of descriptions as possible, all correlated and synchronised) together with the concept of multi-instance support (i.e. as many different media objects as possible for each layer) provide rich and flexible means for encoding music in all its aspects.

It is possible to adopt some ad hoc encoding in addition to already existing formats to represent information. In fact, while a comprehensive format to represent music is not available, popular existing standards must be taken into account. This is not a contradiction because of the two-sided approach of IEEE 1599 to music representation, which is: keep intrinsic music descriptions inside of the IEEE 1599 document – in XML format – and media objects outside of the IEEE 1599 file – in their original format. From a logical point of view, it is possible to organize the contents of a single IEEE 1599 document in two blocks:

1. the group constituted by the General, the Structural, and the Logic layers, whose contents are fully encoded in XML format, and
2. the group including the Notational (e.g., GIF, JPEG, TIFF for a score), the Audio (e.g., AAC, MP3, WAV), and the Performance (e.g., Csound, MIDI, MPEG) layers, that link to external digital objects.

Intrinsic music descriptions, typically catalogue metadata and logical representations of music events, clearly reside inside the IEEE 1599 file, whereas media files maintain their original format and are simply linked from the corresponding layers of the IEEE 1599 file.

Consider the following examples. The symbols that belong to the score, such as chords and rests, are described in XML, in the Logic layer. On the contrary, MP3 files and other audio descriptions are not translated into XML format, rather they are linked and mapped inside the corresponding IEEE 1599 layer, the Audio layer.

It should be clear that the description provided by an IEEE 1599 file is flexible and rich, both in regard to the number and to the type of media involved. Thanks to this approach, a single IEEE 1599 document can contain one or more descriptions of the same music piece in each layer. For example, in the case of an operatic aria, the IEEE 1599 document could house: the catalogue metadata about the piece, its author(s) and genre; the corresponding portion of the libretto; scans of the original manuscript and of a number of printed scores; several audio files containing different performances; related iconographic contents, such as sketches, on-stage photographs, and playbills. Thanks to the heterogeneous information provided by IEEE 1599, software applications based on such a format allow an integrated fruition of music in all its aspects.

The multi-layer approach of IEEE 1599 format is relevant for our work as layers are useful in a shared environment to identify the music piece, to view its score (both reconstructed from symbols and digitized from traditional printed versions), and to listen to the corresponding audio (through both recorded and computer-driven performances). All this information can be embedded within a unique document, so it is easy to retrieve and parse even in a Web-based framework.

After explaining the multi-layer structure of the format, the second key concept to introduce is the spine. In IEEE 1599, the spine consists of a sorted list of events, where the definition and granularity of events can be chosen by the author of the encoding. Due to its role, the spine is a sub-element of the Logic layer.

The spine has a fundamental theoretical importance within the format. It represents an abstraction level, as the events identified in it do not have to correspond to score symbols, or audio samples, or anything else. It is the author of the document who can decide, from time to time, what goes under the definition of music event, according to the needs. Since the spine simply lists events to provide a unique label for them, the mere presence of an event in the spine has no semantic meaning. As a consequence, what is listed in the spine structure must have a counterpart in some layer, otherwise the event would not be
defined and its presence in the list (and in the IEEE 1599 document) would be absolutely useless. For example, in a piece made of n music events, the spine would list n entries without defining them from any point of view. If each event has a logic definition – namely, it is a note or a rest – it is graphically represented in many scores and it is present in a number of audio tracks. These aspects are treated in the Logic, Notational, and Audio layers respectively.

Music events are not only listed in the spine, but also marked by a unique identifier. These identifiers are referred to by all instances of the corresponding event representations in other layers. Thus, each spine event can be described:

- in 1 to n layers; e.g., in the Logic, Notational, and Audio layers;
- in 1 to n instances within the same layer; e.g., in three different audio tracks mapped in the Audio layer;
- in 1 to n occurrences within the same instance; e.g., the notes in a song refrain that is performed 4 times (thus the same spine events are mapped 4 times in the Audio layer, at different timings).

Thanks to the spine, IEEE 1599 is not a simple container for heterogeneous media descriptions related to a unique music piece. It shows instead that those descriptions can also present a number of references to a common structure. This aspect creates synchronization among instances within a layer (intra-layer synchronization), and – when applied to a complex file – also synchronization among contents disposed in many layers (inter-layer synchronization).

The framework we propose is based on a central repository constituted by a set of IEEE 1599 files, whose contents include both descriptions of scores and media objects. As explained before, all the data and metadata related to a single music piece are either embedded into one document or linked to it. We propose to consider a large repository of interacting IEEE 1599 files, in order to provide a very rich environment for Music Information Retrieval.

The central repository is a sort of relational database where i) entities are represented – at different degrees of abstraction – by IEEE 1599 files, layers, and music events respectively, and ii) relationships can take place among IEEE 1599 files, layers, and music events. For instance, the query “Return all the pieces composed by Maurice Ravel” is typically resolved by involving many different files; the query “Extract all the audio tracks corresponding to performances of Ravel’s Bolero” requires the analysis of a single layer (in this case, the Audio layer) within a single IEEE 1599 file; finally, the query “Calculate the maximum voice extension of the flute part from Bolero” addresses the music events described in the Logic layer. Please note that a single IEEE 1599 file is, in a certain sense, a database itself, with its entities and internal relationships.

In conclusion, music-oriented databases could host IEEE 1599 representations of music pieces and the corresponding Web interfaces could integrate applets to provide interaction with music contents such those proposed in [5] and shown in Figure 4.
5. Conclusions

This paper has shown a number of applications concerning music, digital archives, and Web interfaces. Even narrowing the field of interest to single music-oriented databases, some efforts can be done to implement new navigation paths and interaction with metadata. Such advanced models must take into account the intrinsic characteristics of music and music-related materials, in order to provide a wide and well-organized amount of information to common users as well as to experts.

As a further development of Web interfaces to access multimedia archives, we have proposed the integration of IEEE 1599 standard. In this way, the user not only can navigate metadata but also deeply interact with music contents.

6. Bibliography


