Optical Music Recognition System within a Large-Scale Digitization Project

G. Sayeed Choudhury, Tim DiLauro, Brian Harrington

Digital Knowledge Center
Milton S. Eisenhower Library
Johns Hopkins University

Michael Droettboom, Ichiro Fujinaga, Karl Macmillan

Peabody Conservatory of Music
Johns Hopkins University
ich@peabody.jhu.edu

Abstract

An adaptive optical music recognition system is being developed as part of an experiment in creating a comprehensive framework of tools to manage the workflow of large-scale digitization projects. This framework will support the path from physical object and/or digitized material into a digital library repository, and offer effective tools for incorporating metadata and perusing the content of the resulting multimedia objects.

The project involves digitization of the Lester S. Levy Collection of Sheet Music (Milton S. Eisenhower Library, Johns Hopkins University). In Phase One, images of the music and lyrics, and color images of the covers of the Levy Collection were digitized and a database of text index records was created. Phase Two consists of converting the digitized music to computer-readable music notation format along with full-text lyrics, generating sound renditions, and creating metadata to enhance search capabilities.

During Phase One, the researchers at the Eisenhower Library created a database of text index records, images of the music and lyrics and color images of the cover sheets from the Levy Collection. This database is available to the general public at `http://levysheetmusic.mse.jhu.edu`. Currently, the Collection can be searched in three modes. First, users can search by subject, a keyword search on the text record. Each of the pieces has been indexed for the subject of the song and/or cover image. Users may also browse the Collection by the topical arrangement of the physical collection.

In Phase Two, an adaptive optical music recognition (AOMR) software (Fujinaga 1997) is used to convert the TIFF image of scanned sheet music into computer readable-formats, which includes GUIDO and MIDI files along with full-text of the lyrics. These digital objects will be deposited into the data repository along with the scanned sheet music TIFF, JPEG and thumbnail, and associated metadata.

The AOMR software offers five important advantages over similar commercial offerings. First, it can be run in batch processing mode, an essential feature for the Levy Collection given its large number of music sheets. It is important to note that most commercial software is intended for the casual user and does not scale for a large number of objects. Second, the software is written in C and therefore is portable across platforms. Third, the software can “learn” to recognize different music symbols—an issue considering the diversity of the Levy Collection and the universe of notated music, in general. Fourth, the software is open-sourced. Finally, this software can separate full-text lyrics that can be further processed using optical character recognition (OCR) technology.

The AOMR process is divided into two major sections: symbol classification and musical semantic interpretation. The first step in the interpretation phase is to connect all inter-related symbols. In addition, many rhythmic errors can also be corrected by adjusting the metric placement of notes relative to their vertical alignment with notes in other parts.

An interactive graphic editor suitable to be interfaced with the AOMR program is being developed jointly with the group working on the GUIDO editor (Renz 2000). The purpose of this editor is to correct any errors generated by the AOMR so that the corrected version then can be converted to GUIDO format.

To enable powerful search and retrieval as well as user-friendly navigational mechanism, Phase Two of the Levy Project will include a strong metadata component. Commonly defined as "data about data," metadata is structured representational information. The kinds of metadata important for Levy include descriptive (to enable searching, browsing and identification of items), structural (to enable the creation of an interface for optimum browsing and navigation), and administrative (to manage the digital components of the collection and aid users in identification of items).

To further enhance the scholarly value of the Levy Collection, a web interface will be developed for a music research toolkit, for example, Humdrum (Huron 1997). These toolkits are software tools intended to assist in music research and are suitable for use in a wide variety of computer-based musical investigations, such as motivic, stylistic, and melodic analysis and concordance studies. We also propose to extend plans for developing automated means of mining authoritative name information and creating even richer name indexes.

The entire project is an experiment in developing a comprehensive framework of tools to manage the workflow of large-scale digitization projects. This framework will support the path from physical object and/or digitized material into a digital library repository, and offer effective tools for incorporating metadata and perusing the content of the resulting multimedia objects. The Levy Collection, with its large size and availability in digital format, is an ideal subject for development and evaluation of this proposed framework.

Acknowledgements

This project is funded in part by the National Science Foundation, the Institute for Museum and Library Services, and the Levy family.

Suggested Readings

Droettboom, M. 2000. Design and implementation of an optical music interpreter. Technical report, Peabody Conservatory of the Johns Hopkins University. `http://www.peabody.jhu.edu/~mdboom/omi/`

Fujinaga, I. 1997. Adaptive optical music recognition. Ph. D. dissertation, McGill University.

Hoos, H. H. and K. Hamel. 1997. GUIDO music notation version 1.0: Specification Part I, basic GUIDO. Technical Report TI 20/97, Technische Universitat Darmstadt. `http://www.informatik.darmstadt.de/AFS/GUIDO/docu/spec1.htm`

Huron, D. 1997. Humdrum and Kern: selective feature encoding. In *Beyond MIDI: The Handbook of Musical Codes*, ed. E. Selfridge-Field, 375–401. Cambridge, MA: MIT Press.

Renz, K. 2000. Design and implementation of a platform independent GUIDO notation engine. In *International Computer Music Conference.* 469–72.