Jose Maceda Collection Digitization at University of the Philippines Center for Ethnomusicology: Towards a Digital Audio Library

Li Jia, Ph.D College of Music, University of the Philippines
As a spearhead force in music research, especially in the area of South East Asia region, the UP Center for Ethnomusicology (UPCE) caters for a gigantic collection of audio materials which covers different musics and musical traditions in the Philippines, South East Asia and representative areas from other continents.
• As an outcome of its former appellation, the “UP Ethnomusicology Archives”, the UPCE hosts an ethnomusicological collection of about 2500 hours of recorded music in open reel and cassette tape formats, under the authorship of Jose Maceda, whose visionary work of putting together these valuable recorded materials left a treasure for ethnomusicology scholarship and research.
The Jose Maceda Collection Digitization Project (UJMCDCP) is a pioneering endeavor at UPCE to digitize the audio recordings, field notes and photographs of the entire collection using modern digital technology, which helps prolong the limited life span of previous venerable medias to provide access over long period of time.
Objectives

• To make restorative and preventive conservation of the audio materials recorded on analogue medias through modern digital technology.
• To establish a new catalogue and systematic musicological database;
• To provide a title/song-based song indexing/search engine
• To provide enhanced services of access towards a digital audio library which transcends the limitation of time, location and data storage deterioration.
From November, 2009, the project was able to digitize around 1530 open reels of Filipino music and 800 open reels of international sources, as well as 191 cassette tapes, which equates to around 2500 running hours of music. As of August 2012, the catalogue has already accumulated 2649 titles, with the steadfast expansion as the project is still going on.
Audio digitization working flow chart

- Content Digitization
- Content Editing
- Content Storage
- Content Cataloguing
- Content Distribution
1. Content Digitization
The digitization audio workstation has the capability as follows: CPU dual core 2.4 GHZ, Memory – 2 GB, DDR2, two hard disk drives (80 GB + 150 GB), with Windows XP as operating system. The audio interface is RME fireface 400. For the signal input equipment, a Studer A807 MKII function as open reel player and Tascam CD – A750 as cassette tape playback deck.

The major software in the digitization is Wavelab 6 by Steinberg, together with two signal and spectrum analyzing software, Digicheck (bundled tree from the RME interface) and Scan Input.
Slicing tape replacing
The different speeds that the music was recorded produced different specifications of frequency response at record/reproduce mode. Table 1 shows the most frequently encountered speed ratio during this project and their corresponding frequency response.
Table 1

| Speed ratio | 3.75 ips | 7.5 ips | 15 ips |

| Frequency   | +/- 2 dB   | 30 Hz ~ 12 KHz | 30 Hz ~ 16 KHz | 30 Hz ~ 20 KHz |
| Response    | +/- 1 dB   | 30 Hz ~ 8 KHz  | 30 Hz ~ 12 KHz | 30 Hz ~ 18 KHz |
Table 2. Approximating playing times for different reel sizes (one side)

<table>
<thead>
<tr>
<th>Type of Reel/Recording Speed</th>
<th>5” Reel 600ft</th>
<th>7” Reel 1200ft</th>
<th>10.5” Reel 2500 ft</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 7/8 ips</td>
<td>1 hour</td>
<td>2 hours</td>
<td>4 hours</td>
</tr>
<tr>
<td>3 3/4 ips</td>
<td>30 minutes</td>
<td>1 hour</td>
<td>2 hours</td>
</tr>
<tr>
<td>7 1/2 ips</td>
<td>15 minutes</td>
<td>30 minutes</td>
<td>1 hour</td>
</tr>
<tr>
<td>15 ips</td>
<td></td>
<td></td>
<td>30 minutes</td>
</tr>
</tbody>
</table>
Diticheck singal analyzing interface
Scan Input spectrum analyzing interface
Wavelab 6 main interface
2. Content Editing

- Due to the nature of analogue recording of open reels and cassettes, one side of open reel provides a variety length of playing time, as previously shown in Table 2, depending on the length of the tape and recording speed ratio. The same case is with cassette tapes. Each side consists of a various numbers of songs by the same or different researchers. To serve the purpose of searching, indexing and appreciation purpose, the next step was to edit the audio file and cut it by the unit of song/title with corresponding metadata.
• This stage was completed at a separates editing workstation, transfer by external hard disk saved in mp3 format. The specifications of the editing workstation are: dual core 3.0 GHz CPU, 2GB DDRII memory with Linux/windows XP dual operating system. The audio editing software was Audacity 1.312 – beta, which is a free software available for download. The advantage of the software is that it operates on a copy of the original file instead of physical changing realtime, which provides high data security.
Audacity 1.312 – beta file opening window
audacity 1.312 – beta speed change window
Secondly the context cutting and slicing demands a high expertise in music understanding, to assure the cutting point is at the right place. This expertise differentiates the digitization of audio materials towards a digital audio library from other digitization projects. On top of the fact that all the members of the team had the professional training and background in music research they actually listened to all the music, corresponding to the original data from the recordings, to locate exactly where the song stopped, rather than mechanically relying on wave spectrum which is unsecure.
Audacity 1.312 – beta file export window
3. Content Storage
Western Digital mirror edition Caviar®
Green™
• The three versions of digital audio data are stored in separate sets of hard disk. The archival version will be at UPCE for storage. The researcher version will be forwarded to UPCE library for any scholarly research, which can be easily uploaded to a computer or digital CD or DVD. The online version will be for website uploading which would be still a process to follow in the future.
• A. Archival version. The digitized original files are saved in 96KHZ/ 24bit wave format audio file format for archival purpose.

• B. The research version from the digital editing was saved in MPE G -2 Audio Layer III (MP3) format.

• The web version from the digital editing is saved in ogg format.
### 4. Content Cataloguing

<table>
<thead>
<tr>
<th>Identification</th>
<th>Format</th>
<th>Locator</th>
<th>Description</th>
<th>Cataloguer/ Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPCE-A-#</td>
<td>acetate or polyester reel, tape, cd</td>
<td>wav/mp3</td>
<td>all descriptions/metadata available</td>
<td></td>
</tr>
<tr>
<td>Ayta 1971 Logan P1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UPCE-A-0001</td>
<td>1/4&quot; Open Reel</td>
<td>wav/mp3</td>
<td>Sasindog ka koncial &quot;Konsehal&quot;; vocal, instrumental; female (singer); kudlong and saluoy (instruments); both plucked lute (Chordophone)</td>
<td>Opiso / June 3, 2010</td>
</tr>
<tr>
<td>Ata 1971 Side A</td>
<td></td>
<td>A1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the new digital catalogue consist of such parameters as: data identification number (title accession number), old catalogue number, original format, digital format, location of the original data, electronic location, researcher, year, ethnic group, country of origin, meta data description, etc. This provided an efficient system to search online or at UPCE library by providing key words.
5. Content Distribution

- a website hosted at UPCE website (www.upcenom.com) is exclusively created for this digital audio collection. A search engine was design to access by accession number, researcher, ethnical group, country of origin or year of recording.
### Advanced Search

#### Search Controls
- **Search**: [Input Field] [Dropdown] [Search Button]
- **Advanced Search**: [Link]

#### Advanced Search Fields
- **Type**: [Dropdown]
- **Accession No.**: [Input Field]
- **Old Catalog No.**: [Input Field]
- **Researcher First Name**: [Input Field]
- **Researcher Last Name**: [Input Field]
- **Group/Country**: [Input Field]
- **Year**: [Input Field]
- **Search Button**

#### Search Results Table

<table>
<thead>
<tr>
<th>Accession No.</th>
<th>Researcher Name</th>
<th>Collection Type</th>
<th>Group/Country</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPCE-A-0001</td>
<td>Isaias Lim</td>
<td>Music</td>
<td>Ata / Philippines</td>
<td></td>
</tr>
</tbody>
</table>
Conclusion

• Digitization is an evolving process along side with the advancement of digital information technology. In this sense, the ongoing Jose Maceda Digital Audio Library Project is not an end in its own term. It opens a new door towards a institutionalizing digital audio library with a mutual aims towards enhancing scholarly research, share musicological database and promote different indigenous or regional musical traditions and cultural identities.
Concerns

• Firstly, this initial project on digitizing Jose Maceda Collection aims to preserve the original analogue material which demands authenticity of digitized copy as close to the original sources. However, how to further enhance the current digital audio data with the aid of modern digital audio technology, such as noise elimination, post-production effect, etc, will be of vital importance to extend the utility of current digital data to the grassroots of the society for the purpose of music appreciation. This extension of the digitization audio projects can further produce audio data at the fidelity ratio sufficient for publication in CD or DVD formats which can reach a wider audience of households.
Secondly, the Jose Maceda collection also consists of a vast amount of related ethnomusicological materials such as musical transcriptions, transcribed texts, photos, field notes, etc. How to incorporate these data as metadata into the digital audio library is an important topic to solve to boost up the serviceability of the current digital audio library to a higher level in the future.
Lastly, how to establish a more efficient music retrieval system which can describe the current audio data with more musicological data will be of vital importance for future scholarly research. New advancements in digitization like music information retrieval system, Melody Indexing System (MELDEX), adaptive content-based music retrieval system have already paved way for this purpose, from a technological perspective in general. However, how to install an efficient system which serves best the music of indigenous culture and record at old analogue technology live with low-fidelity is still requires for the collaboration and endeavor among musicologist, library scientists and engineering scientists. \textit{In this perspective, establishing of digital audio library is not a simple technical issue, but rather an integral part of ethnomusicology research that has been providing new knowledge and data to support and enhance this area of scholarship.}