

Digital video preservation & access: Broad strokes to local implementation

Amanda Focke, afocke@rice.edu

Monica Rivero, monica.p.rivero@rice.edu

Jane Zhao, janezhao@rice.edu

Fondren Library, Rice University

I. Introduction

Fondren Library worked in 2012-2013 to make a collection of born-digital video oral histories, a part of Special Collections and Archives, available through its institutional repository. The process required research at every step and we had to figure out 1). what guidelines to use to direct our project; 2). what video formats, audio and video codecs to use for archival masters and streaming access copies; 3). What tools to use for converting video files; 4). How to stream access copies; 5). How to best describe the videos and make them findable.

II. Our Guidelines: Broad Strokes/Research Highlights

We looked at George Blood's 2011 white paper, *Refining Conversion Contract Specifications: Determining Suitable Digital Video Formats for Medium-Term Storage*, as guidelines on the nature of source materials, options for masters and access files. The paper is unpublished, but presented at the 2011 Conference of the Association of Moving Image Archivists, Austin, TX. Refer the following link to download and see the full document. <http://www.georgeblood.com/Resources.html>. Specifically, we use the information on page 78, which is *Category 4, Born Digital, Not-tape Formats*, to direct our project. The bulleted points on the slide are, *First, remove from carrier! Second, evaluate for format obsolescence. Third, use native bit depth and encoding, native raster as masters or 8- or 10-bit uncompressed, native raster.*

III. Local Implementation Details: <http://bit.ly/HmI7GG>

1. Analysis of Source Files


The source files include 313 MTS files with H264 video codec and AC3 audio codec, 107 MPG files with MPEG2 video codec and MPEG2 audio codec, and 3 MP4 files with H264 video codec and AAC audio codec.

2. Determination of Archival Masters and Streaming Access Copies

We deem the source files of MPG and MP4 are current and we will support them. We use them as the archival masters without any change. We deem the file wrapper MTS is less compatible with media players and AC3 is a proprietary audio codec, we are not going to support these two technologies.

However, H264 is current and well adopted. So for the source files of MTS, we keep the video codec, and convert the audio codec to AAC and change the file wrapper to MOV.

For the access streaming copies, we decided to use MP4 file with H264 video codec and AAC audio codec, which is well adopted and supported currently.

| # of Video Files | File Wrapper | Video Codec | Audio Codec |  | Masters | Access Copies |
|------------------|--------------|-------------|-------------|---|--------------------------------|------------------|
| 313 | MTS | H264 | AC3 | | MOV + H264 + AAC | MP4 + H264 + AAC |
| 107 | MPG | MPEG2 | MPEG2 | | Source files serve as masters. | |
| 3 | MP4 | H264 | AAC | | | |

3. Tools Used for Converting Files

Open source free video tools FFmpeg, ffmpeg and HandBrake are used throughout the project to create masters and streaming access copies. All the three tools allow users to batch convert files, which is a must-have feature for processing large video collections. Local recipes were created for fine tuned conversions. The documentation on recipes and steps is available at <https://digitalriceprojects.pbworks.com/w/page/67672569/VAHF-master-access-recipe-steps>

4. Streaming Access Copies

HTML5 video streaming without a streaming server doesn't allow viewers to skip forward and backward. The final resort is that we piggyback on Rice Central IT's Flash Streaming Server for a smooth playback of videos and embed JW player into our codes so that the video playback doesn't depend on players on end user's computer.

5. Metadata Challenges

Currently, we only have limited metadata available for the videos. The technical data including duration and a poster image extracted from the original file have been added to the existing metadata. However, we desire to transcribe the videos, and link them to video timecode and embed descriptive metadata into video files.

IV Final Project Online

This special collection of Vietnamese American Heritage Foundation (VAHF) interviews were added to the existing Houston Asian American Archives (HAAA) oral histories that was created and made available in 2010. The collection is available at <http://scholarship.rice.edu/handle/1911/36136>