Feature

Recommended best practices for digital image capture of musical scores

Jenn Riley and Ichiro Fujinaga

The authors

Jenn Riley is Digital Media Specialist at Digital Library Program, Indiana University, Bloomington, Indiana, USA. **Ichiro Fujinaga** is Assistant Professor of Music Technology in the Faculty of Music, McGill University, Montréal, Québec, Canada.

Keywords

Music, Digital documents, File structures, Imaging, Data capture

Abstract

Like other complex visual articles with small details, musical scores are difficult to capture and present well in digital form. This article presents methods that can be used to reproduce detail and tone from printed scores for creating archival images, based on best practices commonly used by the library community. Capture decisions should be made with a clear idea of the purpose of the imaging project yet be flexible enough to fulfill unanticipated future uses. Options and recommendations for file formats for archival storage, Web delivery and printing of musical materials are discussed.

Electronic access

The Emerald Research Register for this journal is available at
http://www.emeraldinsight.com/researchregister

The current issue and full text archive of this journal is available at
http://www.emeraldinsight.com/1065-075X.htm

OCLC Systems & Services
Volume 19 · Number 2 · 2003 · pp. 62-69
 · ISSN 1065-075X
DOI 10.1108/10650750310481784

Introduction

Libraries and archives embarking on digital imaging projects today have more guidance for decision making than just a few years ago. Standards and best practices for digitizing various types of originals have emerged, from the early NARA (National Archives and Records Administration) Guidelines (Puglia and Roginski, 1998), Cornell University's *Digital Imaging for Libraries and Archives* (Kenney and Chapman, 1996), its successor *Moving Theory into Practice* (Kenney and Rieger, 2000), and the Library of Congress's documentation for the American Memory project (Fleischhauer, 1998; Library of Congress, 2000) to the Arts and Humanities Data Service's Guides to Good Practice series (Arts and Humanities Data Service, 2002) and the NINCH (National Initiative for a Networked Cultural Heritage) Guidelines (NINCH, 2002).

Different standards and best practices documents take different approaches to dictating image specifications. Some put forth prescriptive lists of appropriate resolutions and bit depths for various formats, while some describe decision-making processes for determining specifications for digitizing individual items. They tend to focus on photographic and printed textual materials, rather than more specialized materials such as maps and musical scores. Fortunately, many approaches mentioned in these guidelines can be adapted to the digitization of musical scores. However, musical notation, by its very nature, requires accurate capture of small details. Staff lines, ledger lines, dots, and bars must all be adequately captured or the notation suffers significant loss of meaning. This paper will present some best practice guidelines that can be applied to the capture of musical scores in digital format.

Purposes of scanning

Before capture specifications are determined, the purpose of the imaging project must be

Ichiro Fujinaga would like to thank Michael Droettboom, Karl MacMillan, and Asha Srinivasan for their help in the preparation of this paper. This research is funded in part by the NSF's DLI-2 initiative (#981743), IMLS National Leadership Grant, and support from the Levy family.

OCLC Systems & Services
Volume 19 · Number 2 · 2003 · 62-69

clearly defined. Mass-printed editions of common works in poor condition may require an approach that maximizes the capture of detail from the musical notation itself, without attempting to reproduce any other visual elements on the page. However, manuscripts, rare materials, and scores with annotations by important collectors require faithful reproduction of all markings on the page, treating the score as a historical artifact. Digital capture of paper watermarks from music manuscripts (Edge, 2001; Kenney and Rieger, 2000; Stewart *et al.*, 1995; Wenger *et al.*, 1995) would require yet a third approach, which will not be covered in this article. It is important to note that not all materials are good candidates for digitization. For example, rare or fragile materials might best be captured for preservation purposes on medium-format color film, such as Ilford's Ilfochrome, which has an estimated 300-year life expectancy.

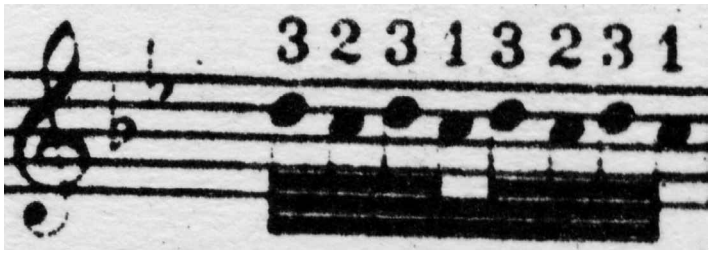
Master file specifications

Although we cannot predict all future uses for digital images, good practice dictates that the capture process create a flexible "master" image, from which images for a wide variety of specific uses can be derived. At the very least, master images of musical scores must support the creation of derivative versions for Web delivery and printing, and contain enough detail for successful Optical Music Recognition (OMR). To ensure this high level of flexibility, four factors must be considered: capture resolution, color reproduction, choice of master file format, and storage of master files.

Resolution

Scanning resolution must be set to capture all important details from the originals. One method used to determine minimum scanning resolution for illustrations involves measuring the width of the smallest stroke and setting the scan resolution to capture that stroke with a designated number of pixels (Kenney and Rieger, 2000, pp. 46-7). For musical notation, this smallest detail is generally the white space between beams (see Figure 1). In the Figure, the beams are thick horizontal black rectangular bars that connect and group notes. While Kenney advocates

Figure 1 An example of very small spacing between beams



capturing the smallest detail with 2 pixels for adequate reproduction of the stroke with a grayscale scan, 3 pixels per detail is required for successful OMR with the forthcoming Gamera software (MacMillan *et al.*, 2001). However, details in musical notation are consistently smaller than 1mm and are difficult to measure accurately without specialized equipment. Also, since print size varies between publications, this method would have to be applied individually to each piece of music to be scanned.

Because of these problems, it would be more appropriate for most projects to simply capture all images at the same resolution. Our tests have found that 600 dots per inch (dpi) is a sufficient resolution to capture all significant details in most musical notation, as seen in Figure 2. The 600dpi scan renders more adequately the ledger line and the sharp sign. This resolution will capture details as small as 0.005in (0.027mm) with the required 3 pixels. Figure 3 shows the same image captured at a resolution of 300dpi.

For larger printed notation, 300dpi may be sufficient. Our preliminary studies show that resolutions above 600dpi generally do not offer much advantage for the purpose of Web viewing, printing, or OMR. This is true even in the case of miniature scores, as shown in Figures 4-7. There is an improvement from 300dpi to 600dpi but there are no significant improvements in 1,200dpi or 1,600dpi scans. Grayscale versions of these sample images at their original sizes and a few other examples can be found at <www.dlib.indiana.edu/~jenlrile/oclc/resolution/>

Color reproduction and bit depth

Musical notation must be captured in grayscale, as 1-bit (bitonal) scanning is generally not adequate to capture all important detail (see Figures 8 and 9). The grayscale scan shows

OCLC Systems & Services
Volume 19 · Number 2 · 2003 · 62-69

Figure 2 Small detail scanned at 600dpi

Figure 3 Small detail scanned at 300dpi (resized for comparison with Figure 2)

more adequately the space between beams, despite the lower resolution. If the page uses color to convey important information, such as the color on sheet music covers, color scanning should be used. Grayscale scans should be at least 8-bit, while color scanning should be at least 24-bit. Higher bit depths may be appropriate for some uses. In order to preserve this full color range, any image manipulations done according to the guidelines below should be performed in the scanning software at the time of capture, not by editing the image after capture. It is important to understand that image manipulations on the master file, such as color correction and straightening, will reduce the amount of data present in the master image, which affects its usefulness.

Before performing any image adjustments, the imaging system must be set up properly to ensure color fidelity throughout the digitizing process. This is done with International Color Consortium (ICC) profiles created through specialized software from companies such as Monaco Systems <www.monacosys.com/index.html> The user provides a target with known color values to the device to be characterized (a scanner, monitor or printer). The profiling software, often with the aid of a colorimeter, then compares the known values with what that device actually registers and creates a mapping between the two. This ensures that the digital image encodes color values that are not connected solely to any particular piece of hardware used for scanning, display, or printing.

Once a system is set up in this manner, it should capture reasonably color-accurate versions of the original printed materials. If the purpose of the imaging project is to capture the artifact as it exists today, no corrections should be made to the master images. Every effort should be made to ensure pages are straight during capture as rotating them in image-editing software can result in a loss of detail. If capture of the musical content rather than visual content has been determined as the purpose of the scan, the contrast between the musical notation and background of the page should be maximized. A well-contrasted page will use the entire tonal range and have completely filled-in note heads, solid staff lines, and clean white space between staff lines when viewed at 100 percent magnification in image-editing software, as seen in Figure 10. Figure 11 shows an example of a poorly-contrasted image.

Master file formats

Uncompressed TIFF is generally suggested as the most appropriate file format for master files (Fleischhauer, 1998; Puglia and Roginski, 1998). However, TIFF is not a true but a *de facto* standard. The extensibility features of TIFF have resulted in the emergence of many variations of the format, and it is proving difficult for developers to effectively support these. PNG, on the other hand, is a more modern and a simpler file format than TIFF, while maintaining the necessary requirements as a master file format. One of the major advantages of PNG is its ability to use patent-free lossless compression, which produces significantly smaller files than uncompressed TIFF and various JPEG lossless compression schemes (Santa-Cruz *et al.*, 2000). Most archival imaging projects, however, still use TIFF for master files, and it may be some time before it is clear whether the digital library community as a whole accepts PNG as a master file format.

Storage of master files

Proper storage of master files is perhaps the most difficult aspect of managing a digital imaging project. Storage of master files on optical media such as CD-R and DVD-R is a short-term solution and should be supplemented by a long-term refreshing policy,

OCLC Systems & Services
Volume 19 · Number 2 · 2003 · 62-69

Figure 4 Miniature score scanned at 300dpi

Figure 5 Miniature score scanned at 600dpi

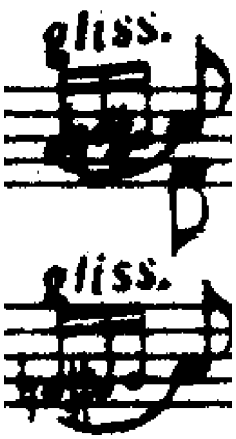


Figure 6 Miniature score scanned at 1,200dpi

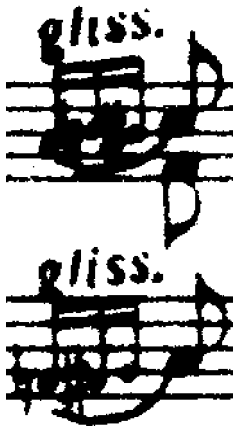
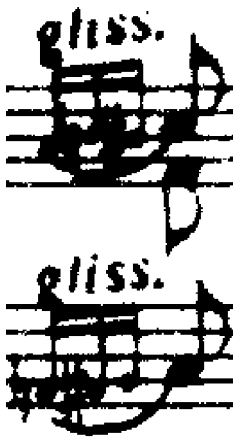


Figure 7 Miniature score scanned at 1,600dpi



so that data can be digitally transferred to newer storage media before the older media become obsolete. One possible system allowing for multiple copies of master and derivative files on a variety of media may consist of one or more servers with hard disks used to store both master and derivative files for online access with a tape back-up system in case of server disk failure.

Additionally, three copies of master files are stored on DVD-Rs, and two of those should be stored off-site. A specific list of hardware and media involved in such a system is included in Appendix 1. Although the cost of hardware and media is decreasing, it is important to stress the need for continuing technical support for server maintenance, back-ups, and refreshing when embarking on digital projects.

Web delivery

File formats

At first glance, there appear to be a large number of file format options for Web delivery. Ideally, a file format for Web delivery of musical scores will meet three criteria, listed in order of importance:

(1) be viewable by the target user population with their preferred Web browser and installed plug-ins;
(2) yield file sizes appropriate for delivery to target users over their network connections, which may be quite slow; and
(3) support multi-page images that package entire scores into single files.

Table I shows some possible file formats for Web delivery and their support for the first and third criteria.

OCLC Systems & Services
Volume 19 · Number 2 · 2003 · 62-69

Figure 8 Example of a grayscale scan at 600dpi

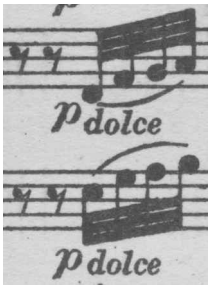
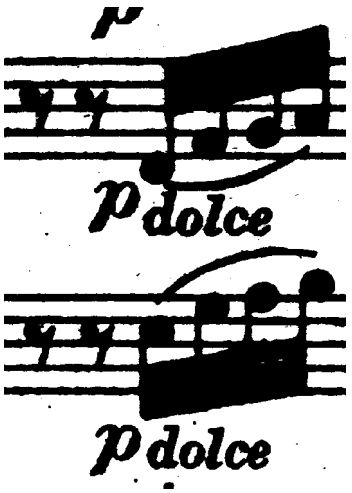


Figure 9 Example of a bitonal scan at 1,000dpi



Unfortunately, there is no single file format today that meets all three of these criteria, so some trade-offs must be made. The first and most important criterion listed for determining Web-deliverable file formats is viewability. Although DjVu (Bottou *et al.*, 1988) and JPEG2000 (Santa Cruz *et al.*, 2000) meet the other criteria for Web delivery extremely well, the lack of widespread viewing software seriously limits their usefulness at this point in time. Formats that are not viewable with commonly-used software should be used sparingly, either in situations where images are restricted to a certain audience with known system configurations, or as an alternative to a more accessible format for the general public.

Figure 10 Well-contrasted image

Figure 11 Poorly-contrasted image

Table I Comparison of some Web-deliverable image formats

File format	Commonly viewable via the Web	Multi-page support
JPEG	✓	
GIF	✓	
PNG	✓	
TIFF		✓
PDF	✓–	✓
DjVu		✓
JPEG2000		✓

Among these commonly viewable formats, PNG has a reputation for poor Web browser support. However, only advanced functionality such as alpha transparency, which allows for a variable level of transparency for each pixel, has support problems in current browsers. Score images do not need this advanced functionality and thus PNG is an acceptable choice for Web delivery of score image files.

File size is the second criterion for a Web delivery format. PDF may at first appear to be an attractive option for score Web delivery because the viewing software is fairly pervasive. PDF also supports multiple pages in a single file, and the Web-viewable text files tend to be extremely small. However, PDF files created from images at acceptable resolutions for screen viewing and printing, even for short pieces, generally have prohibitively large file sizes. Therefore, only formats that do not meet the third criterion, support for multiple pages in a single file, are left for serious consideration, including JPEG, GIF, and PNG. Since the formats themselves do not link pages to one another, a "page-turning" mechanism must be built into the user interface. The choice between JPEG, GIF, and PNG is related to the pixel dimensions of the delivered images, as described in the next section.

Pixel dimensions

The size of a score image for screen display depends on the size and type of the original, and the needs of users. Most standards and best

practices for digital image Web delivery give recommendations on pixel dimensions necessary for displaying the most image detail possible, while ensuring the entire image fits in a Web browser window. However, for musical notation, the readability of the page and the level of detail presented are essential, and thus are more important than making an entire score page visible at a glance.

Down-sampling master image files to 100-200 dots per inch (of the original page size) should result in screen-readable images for most originals. Some sample images for Web delivery illustrating the level of detail that can be displayed can be found at <www.dlib.indiana.edu/~jenlrile/oclc/webdelivery/> A screen shot of some of these images is available in Appendix 2. As the sample image pixel sizes show in Table II, larger originals will need down-sampling to lower resolutions near 100dpi, and smaller originals can be down-sampled at a lower rate to resolutions near 200dpi to show all necessary detail and fit on the screen without horizontal scrolling. However, vertical scrolling will be required in most cases.

At these sizes, there is very little visual difference between grayscale JPEG, GIF, and PNG files of musical score pages. JPEG files are preferable to GIF files for two reasons. We have found that, for grayscale notation pages, JPEG images of score pages at medium-high to high quality tend to be smaller than GIF files, and do not show significant compression artifacts at these sizes. Scores with large printing can be compressed more heavily, down to what many define as "medium" quality (e.g. 50 percent in utilities such as ImageMagick and GraphicConverter or level 6 in Adobe Photoshop). For color images, GIF files are unsuitable because the GIF format is limited to an 8-bit palette, which can result in unacceptable color shifting. PNG offers an advantage over JPEG in that it can use lossless compression. We have found that PNG files for delivering score images on the Web are smaller than high-quality JPEGs but larger than medium-high quality JPEGs. Some average file sizes for the different formats can be found in Table III.

Table III Representative file sizes for Web-deliverable images from 9" × 12" original

File format	200dpi	150dpi	100dpi
GIF	598K	389K	216K
PNG	500K	326K	180K
JPEG high quality	647K	421K	280K
JPEG medium-high quality	411K	268K	137K
JPEG medium quality	332K	215K	111K

For some collections it may be appropriate to provide thumbnail-sized images for browsing, but thumbnail-sized images of music notation are generally not very useful because the notation cannot be read. However, for some collections, such as illustrated sheet music covers, thumbnails may be an effective means of browsing. These may be created by down-sampling master files to 5-25 dots per inch (of original page size). The compression method can be either medium to high quality JPEG or PNG.

Printing

Users of digitized musical score collections have a greater need for printable versions than many other types of originals. While it may not be important to be able to print colored covers or pages from original manuscripts, score pages intended for practice or performance use must be printable. While the absolute best file format for

Table II Sizes of Web-deliverable images

	200dpi	150dpi	100dpi
5.5" × 7.5" miniature score	1,100 px × 1,500 px will not fit horizontally on many common screen resolutions	825 px × 1,125 px adequate for most purposes, but still requires horizontal scrolling for smaller screen resolutions	550 px × 750 px will fit horizontally on all common screen resolutions
9" × 12" score or sheet music	1,800 px × 2,400 px will not fit horizontally on any common screen resolution	1,350 px × 1,800 px requires horizontal scrolling for most common screen resolutions	900 px × 1,200 px will fit horizontally on all but the smallest common screen resolutions

OCLC Systems & Services
Volume 19 · Number 2 · 2003 · 62-69

print versions of score images may vary among user populations, generally score images for printing on laser printers are best presented as bitonal files at 250-400dpi, depending on the original print size (see examples at <www.dlib.indiana.edu/~jenlrile/oclc/printing/>). At lower resolutions, bitonal PNG files on average are smaller, while at higher resolutions Group 4 compressed TIFF files on average are smaller, as shown in Table IV.

Files intended for printing must be easily downloaded by users. The TIFF format allows multi-page files, which would eliminate the need for bundling single image files using a utility like ZIP for Windows or TAR for Unix-based systems. However, many TIFF viewers cannot display multi-page TIFF files.

Image processing for derivative creation

While no image processing should be done on master files, it may be appropriate to perform image manipulations when creating derivative images for specific uses. Resizing to dimensions appropriate for Web delivery or printing is the most common image-processing step when creating derivatives. Other processing steps needed depend on the image's intended use. Web-deliverable versions may require sharpening or color correction to be more readable on-screen. Images for printing may require thresholding (for conversion from grayscale to bitonal) and deskewing (straightening). The exact image-processing steps needed in order to produce the best-quality derivative for a specific use depends on the characteristics of the master file. The planning stages of a digital imaging project must include an evaluation of the derivatives needed and testing of the processing steps needed to create them.

Table IV Bitonal image file size comparison between PNG and Group 4 compressed TIFF file formats

dpi	PNG	TIFF (Group 4)
800	329KB	192KB
400	183KB	146KB
250	90KB	96KB
200	64KB	71KB
100	25KB	38KB

Conclusion

Digital imaging standards and best practices can be applied to the digitization of musical scores, when used with a full understanding of the decision-making processes behind their recommendations. A well-designed digital imaging process with appropriate quality control mechanisms can result in flexible master files from which successful OMR can be done, and Web-viewable and print-quality images can be created.

References

Arts and Humanities Data Service (2002), "Guides to good practice in the creation and use of digital resources", available at: www.ahds.ac.uk/guides.htm (accessed February 12 2003).

Bottou, L., Haffner, P., Howard, P.G., Simard, P., Bengio, Y. and Le Cun, Y. (1988), "High quality document image compression with DjVu", *Journal of Electronic Imaging*, Vol. 7 No. 3, pp. 410-25.

Edge, D. (2001), "The digital imaging of watermarks", *Computing in Musicology*, Vol. 12, pp. 261-74.

Fleischhauer, C. (1998), "Digital formats for content reproductions", *Library of Congress*, July 13, available at: http://memory.loc.gov/ammem/formats.html (accessed February 12 2003).

Kenney, A. and Chapman, S. (1996), *Digital Imaging for Libraries and Archives*, Cornell University Library, Ithaca, NY.

Kenney, A. and Rieger, O. (2000), *Moving Theory into Practice*, Research Libraries Group, Mountain View, CA.

Library of Congress (2000), "Building digital collections: technical information and background papers", April 14, available at: http://memory.loc.gov/ammem/ftpfiles.html (accessed February 12 2003).

MacMillan, K., Droettboom, M. and Fujinaga, I. (2001), "Gamera: a structured document recognition application development environment", *Proceedings of the 2nd Annual International Symposium on Music Information Retrieval*, October 15-17, Bloomington, IN, pp. 15-16.

NINCH (2002), *NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials*, October, available at: www.nyu.edu/its/humanities/ninchguide/index.html (accessed February 12 2003).

Puglia, S. and Roginski, B. (1998), "NARA guidelines for digitizing archival materials for electronic access", January, National Archives and Records Administration, available at: www.archives.gov/research_room/arc/arc_info/ guidelines_for_digitizing_archival_materials.html (accessed February 12 2003).

Santa-Cruz, D., Ebrahimi, T., Askelöf, J., Larsson, M. and Christopoulos, C.A. (2000), "JPEG 2000 still image coding versus other standards", in *Proceedings of the SPIE's 45th Annual Meeting, Applications of Digital Image Processing XXIII*, Vol. 4115, pp. 446-54.

Stewart, D., Scharf, R.A. and Arney, J.S. (1995), "Techniques for digital image capture of watermarks", *Journal of Imaging Science and Technology*, Vol. 39 No. 3, pp. 261-7.
Wenger, E., Karnaukhov, V., Haidinger, A. and Merzlyakov, N. (1995), "Image analysis for dating of old manuscripts", in Chin, R.T., Ip, H.S., Naiman, A.C. and Pong, T.-C. (Eds), *Image Analysis Applications and Computer Graphics*, Lecture Notes in Computer Science 1024, Springer, Berlin.

Further reading

Roelofs, G. (1999), *PNG: The Definitive Guide*, O'Reilly, Sebastopol, CA.

Appendix 1. Sample system for storage and delivery of digital images

Total storage (3TB) cost (as of February 2003): $49,000:

- Web server with 3TB: $17,000
  - Dell server: $5,100
  - Arena Indy 2600 16bay RAID5 IDE-SCSI disk array: $6,300
  - 16 × 200GB disks ($350 each): $5,600
- Autoloader 4TB Dell PowerVault 1288T LTO tape back-up (200GB Ultrium): $10,000
- 60 Ultirum tapes × $55: $3,300
- Pioneer DRM3000 DVD-burner: $15,000
- 2000 DVD-R × $0.55: $1,100 (three copies of 3TB back-up)
- UPS, extra hard disks, extra media: $2,600

The basic premise of this configuration is the assumption that most of the data on this server will not change much once the digitization process is complete; thus the use of DVDs for long-term storage and relatively small number of tapes for back-up is appropriate. These specifications, however, will constantly change; for example, 400GB Ultrium tapes will be available in the near future.

Appendix 2. Resolution requirements for Web delivery of musical scores

(See Figure A1.)

Figure A1 Resolution requirements for Web delivery of musical scores