Recommended Best Practices for Image Capture of Musical Scores

Ichiro Fujinaga
Peabody Conservatory of Music
Johns Hopkins University
1 E. Mt. Vernon Pl.
Baltimore, MD 21202
+1 (410) 659-4100
ich@jhu.edu

Jenn Riley
Digital Library Program
Indiana University
1320 E. 10th St., E170
Bloomington, IN 47405
+1 (812) 856-5759
jenlrile@indiana.edu

ABSTRACT

The success of information retrieval depends heavily on the quality of data input into them. Musical scores, as a complex visual format with small details, are particularly difficult to digitally capture and deliver well. Virtually all capture decisions should be made with a clear idea of the purpose of the resulting digital images. Master images must be flexible enough to fulfill unanticipated future uses as well. In order to provide a framework for decision-making in musical score capture projects, best practices for detail and color capture are presented for creating an archival image containing all relevant data from the print source, based on commonly defined purposes of digital capture. Options and recommendations for file formats for archival storage, web delivery and printing of musical materials are presented.

1. INTRODUCTION

Music information retrieval research often deals exclusively with methods of resource discovery. However, the quality of the data itself is a critical part of the retrieval system, as content-based retrieval cannot work on inferior content. To analyze erasures on a digital copy of a manuscript page requires accurate detail and color capture. To convert scanned score images to notation via Optical Music Recognition (OMR) requires that the original scan contain enough information to successfully perform the OMR. A system that delivers score images to end-users must provide them with a sufficient level of detail for their purposes, or the retrieval has failed. Capture of musical page images is not a trivial process.

Musical scores are difficult to digitally capture and deliver for several reasons. They contain small details such as staff lines, dots, and bars that are essential to the meaning of the notation. They come in a wide variety of formats (printed and manuscript) and sizes. This document is intended to provide an overview of the decisions that must be made when undertaking a musical score digital imaging project and provide some best practices for making these decisions.

2. DEFINING THE PURPOSE OF SCANNING

The first step in planning any imaging project is to determine the purpose of image capture. Some formats of music, such as original manuscripts, rare imprints, and pages with important pencil markings or annotations, should be captured with the intent of reproducing the current appearance of the page as an artifact. Other formats, such as published, typeset scores now in poor condition, the capture of the current condition of the pages is not an important consideration, but capturing the musical content contained within is. Scanning of manuscripts to capture watermarks in the paper rather than the markings on the paper is yet another possible purpose of capture, but will not be covered here [1][2][3][4]. Some formats, such as printed sheet music with illustrated covers of socio-historical interest, may benefit from a hybrid approach to capture. The guidelines below offer advice on both types of capture: artifactual and content-based.

For overly rare and fragile materials, analog capture on film may be a better option than digital imaging. Generally 4”x5” color film is most suitable for this purpose [5][6].

3. MASTER FILE SPECIFICATIONS

3.1 Resolution

Regardless of the purpose of capture, it is essential that musical notation be scanned at a resolution sufficient to capture all significant detail. For a homogeneous collection, with little variation in print size, it may be appropriate to apply a formula advocated by Anne Kenney of Cornell to determine the minimum scanning resolution based on the stroke width of the smallest detail [7]. For musical notation, we have found this smallest detail is generally the white space between beams. While Kenney advocates capturing the smallest detail with 2 pixels for adequate reproduction of the stroke with a grayscale scan, we have found that 3 pixels per detail is required for successful OMR with the forthcoming Gamera software. [8]

However, this benchmarking approach based on stroke width presents some problems. Details in musical notation are consistently smaller than 1mm, requiring specialized magnifying and measuring equipment to accurately characterize. Also, collections to be scanned are generally not homogeneous, which would require measurement and resolution calculation for each item to be captured. This would add a great deal of additional labor to the capture process.

Because of these problems, for most projects it would be more appropriate to simply capture all images at the same resolution. Our tests have found that 600dpi is a sufficient resolution to capture all significant detail for most musical notation, as seen in Figures 1 and 2, where the 600dpi scan more adequately renders the ledger line and the sharp sign. This resolution will capture detail as small as .005 in (.027mm) with the required 3 pixels. Although the resolution of 600dpi is definitely better than 300dpi, especially with scores with small details, preliminary studies show that going beyond 600dpi does not offer much advantage for the purpose of viewing, printing, or OMR. This is true even in the case of miniature scores with extremely small printing. Before any imaging project is begun, these recommendations should be tested with the material and equipment at hand.

Figure 1. 300dpi scan of small detail.

Figure 2. 600dpi scan of small detail.

Illustrated sheet music covers often contain printed halftone illustrations. 600dpi scans should be sufficient to avoid moiré patterns in images from these illustrations [9], but specific illustrations should be examined to determine if higher resolutions or descreening filters are required.

3.2 Color Reproduction and Bit Depth

Many musical scores are printed in black and white. 8-bit or higher grayscale capture of these pages is most appropriate. Some score pages, sheet music covers, and original manuscripts may be in color, however, and these should be captured in at least 24-bit color. In order to preserve this full color range, any image manipulations done according to the guidelines below should be done in the scanning software at the time of capture, not after capture with an image-editing application. If the purpose of capture is both to represent the page as an artifact and to gain a file appropriate for OMR or other types of automatic processing, 12-, 14-, or 16-bit per channel capture may be needed.

If accurately capturing the materials in their current state is an objective of the scan, it is essential that all devices in the imaging system, including scanner, monitor and printer if printed output is included in the project, be characterized and managed via International Color Consortium (ICC) profiles [10]. Operating-system level color management exists both for Macintosh in the form of ColorSync [11] and for Windows 98, 2000, ME, and XP in the form of Image Color Management (ICM) [12]. Locally-created ICC profiles for each device are preferable to generic profiles for a specific model of scanner, monitor or printer [13]. Scanning for artifactual purposes will involve simply scanning the item in a fully calibrated system with no manual adjustments at all. While this characterization process will not be perfect due to the lack of commercially-available targets on the same media as the musical materials to be scanned, color management with ICC profiles is the best current solution to the problem of accurate image capture.

If capture of the musical content rather than visual content has been determined as the purpose of the scan, scanner and monitor characterization with ICC profiles is preferred to ensure that the visual image the operator is viewing when making color adjustments is accurate. If monitor profiling software is not possible for the project, utilities like Adobe Gamma Loader (installed with Adobe Photoshop) can provide better results than an uncalibrated monitor. Care should be taken when making adjustments to remember that the image on the monitor may not accurately reflect the colors present in the raw data of the image.

For content-only capture, manual or automatic adjustments may be done in order to maximize the contrast between the musical notation and background of the page. Two steps can be taken to maximize this contrast. First, the white point of the page should be set to gray value 255 (0% black) and the black point of the page should be set to gray value 0 (100% black). In some scanning software, this can be automated and in others it must be done manually to the histogram. The second step is to manipulate midtones in the scanning software, using controls that may be labeled brightness, contrast, or gamma. This should be done after setting the white and black points. Experimentation with specific scanning software must be done to find the settings that will maximize the foreground-background contrast. It is essential that when determining what adjustments to make, the effects of the tool used on the image as a whole are understood. A well-contrasted page will have completely filled-in note heads, solid staff lines, and clean white space between staff lines when viewed at 100% magnification in image editing software. Additional image processing, such as deskewing, despeckling and sharpening, when done at scan time, may enhance the readability of some originals.

3.3 Quality Control

A thorough quality control system is essential for large-scale digital imaging projects. Some image features, such as resolution, bit depth and pixel dimensions can be checked automatically with software like ImageMagick [14], while others, such as contrast settings, may require visual inspection. Random visual review should be done on master images and all output formats as well.

3.4 Master File Formats

Imaging best practices to date have almost unanimously advocated uncompressed TIFF as the format for storage of master files [15]. However, the TIFF specification is copyrighted by Adobe, and its status has traditionally been as a de facto rather than a true standard. The PNG format may be an emerging replacement for TIFF for this purpose [16]. PNG has the capabilities to store all relevant information captured according to these guidelines, and as such would be technically suitable for a master file format for musical score images. PNG uses lossless compression, and as such PNG files are significantly smaller than uncompressed TIFF files [17]. However, most archival imaging projects to date still use TIFF as the master file format, 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.

4. WEB DELIVERY

Regardless of the capture philosophy guiding the creation of master files, creation and delivery of derivative files for web viewing is similar. In considering file formats for web access, the choice is limited to JPEG, GIF, PNG, and PDF as these formats are supported by majority of web browsers. Thus formats such as TIFF and DjVu [18][19] are not considered generally appropriate for this purpose. Delivery by single image file formats such as JPEG, GIF, and PNG requires a page-turning mechanism for navigation. While the PDF format includes this functionality, it was not really intended for navigating raster image files, so the file sizes of full musical scores in PDF format are prohibitively large.

For screen display, grayscale or color images at 100-200dpi (based on original page size) are generally sufficient to display all relevant detail. 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 obvious compression artifacts at these sizes. Since the file sizes are similar for JPEG and PNG files of this type, for content-based viewing, PNG is preferred because of the faster decompression of PNG

files [20]. Since sizes of paper originals and types of end-users' systems differ so drastically, it is not possible to develop specifications for web images that for all master files, will display all necessary detail and still fit on everyone's preferred window size. Therefore, we have considered the display of musical detail more important here and as a result, in some cases users may have to scroll to see all of a display image. Resolution of display images is best set to not require horizontal scrolling.

For some applications, thumbnail files may be part of the interface. 5-25dpi files (based on original page size) should produce thumbnail-sized images, although these may not be very useful for notation pages. At these dimensions, grayscale images with medium JPEG compression offer a significant file size savings over PNG, as seen in Table 1.

Table 1. JPEG and PNG file size comparison for grayscale thumbnail images.

	PNG	JPEG
25dpi	23 KB	10 KB
10dpi	4.9 KB	2.3 KB
5dpi	1.6KB	.8KB

5. PRINTING

Printing of score pages is a necessary capability for many musical notation imaging projects. While it may not be important to be able to print colored covers or pages from original manuscripts, score pages intended for use for practice or performance will need print capability. While the exact best format of the print versions of score images may vary between user populations, generally score images for printing on laser printers are best presented as bitonal files at 250-400dpi, depending on the original paper size. At lower resolutions in this range, bitonal PNG files on average are smaller, while at higher resolutions, CCITT Group 4 compressed TIFF files on average are smaller, as shown in Table 2.

Table 2. PNG and Group 4 compressed TIFF file size comparison for bitonal images.

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

6. CONCLUSION

Consistent and useful information retrieval depends heavily on the presence of adequate data in the first place. While many local decisions must be made for any digital imaging project, the overarching philosophy of capturing and presenting to the user all relevant data in the source material should guide these decisions. The procedures outlined here are based on current best practices in archival digital imaging, and have been adapted to suit the unique capture and delivery requirements of musical notation.

7. ACKNOWLEDGMENTS

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.

8. NOTES AND REFERENCES

[1] Watermarks may be obtained by tracings, transmitted light photography, photosensitive paper technique (Dylux); beta radiography, and microfocus radiography. Each of these methods have advantages and disadvantages. The use of image processing combining reflective and transmissive scan may offer inexpensive and relatively accurate results.

[2] Edge, D. 2001. The digital imaging of watermarks. *Computing in Musicology* 12: 261-74.

[3] Stewart, D., R. A. Scharf, and J. S. Arney. 1995. Techniques for digital image capture of watermarks. *Journal of Imaging Science and Technology* 39(3): 261--7.

[4] Wenger, E. et al. 1995. Image analysis for dating of old manuscripts. *Lecture Notes in Computer Science* 1024: 522-3.

[5] Assuming that the size of the original document is 9"x12", the 35mm film must be scanned at over ten times the effective resolution. For example, to achieve the equivalent of 600dpi scan of the original, the 35mm film (36x24mm) must be scanned at over 6000dpi. The 105mm microfiche film (105x148mm) only needs to be scanned at twice the effective resolution.

[6] Ilford's Ilfochrome Micrographic film seems to the only color film suitable for preservation with an estimated life expectancy of over 300 years. <http://www.microcolour.com/mci03.htm>

[7] Kenney, A., and O. Rieger. 2000. *Moving Theory into Practice.* Mountain View, California: Research Libraries Group. 46-47; and at <http://www.library.cornell.edu/preservation/tutorial/conversion/conversion-05.html>.

[8] MacMillan, K., M. Droettboom and I. Fujinaga. 2001. Gamera: A structured document recognition application development environment. Proceedings of the International Symposium on Music Information Retrieval. 15-6.

[9] Kenney and Rieger 2000, 36.

[10] The ICC's home page is at <http://www.color.org/>.

[11] See <http://docs.info.apple.com/article.html?artnum=11155> for an overview of Apple's ColorSync.

[12] See <http://msdn.microsoft.com/library/default.asp?url=/library/en-us/icm/icmstart_5i91.asp> for an overview of Microsoft's Image Color Management (ICM).

[13] Software packages to perform profiling on your own devices are commercially available, such as those from Monaco Systems <http://www.monacosys.com/index.html>.

[14] The ImageMagick home page is at <http://www.imagemagick.org>.

[15] The National Archives and Records Administration <http://www.nara.gov/nara/vision/eap/digguide.pdf> and the Library of Congress <http://memory.loc.gov/ammem/formats.html> have taken the lead in documenting best practices for digital imaging.

[16] As of 21 May 2001, PNG (ISO/IEC 15948) is under publication. PNG files are typically 20% smaller than GIF files.

[17] The storage cost is becoming a less of an issue today, mainly because of the recent reduction in the cost of DVD-Rs ($2.20 each). A DVD can hold about 4.7GB or about $0.50/GB. Kenney estimates about $0.60/page for scanning cost (Kenney and Rieger 2000, 168), or $60,000 for 100,000 pages. A page of music (9"x12") scanned at 24bit 600dpi can be compressed to about 30MB or 3TB for 100,000pages. This can be stored in less than 650 DVDs or about $1500 or 1300 DVDs and $3000 for two copies. Thus, $3000 is only 5% of the total cost $60,000. If there's no color, 8bit grayscale 600dpi scan can be compressed to about 12MB/page. This results in $1200 (two copies) or 2% of $60,000.

[18] Bottou, L., P. Haffner, P. G. Howard, P. Simard, Y. Bengio, and Y. Le Cun. 1988. High quality document image compression with DjVu. *Journal of Electronic Imaging* 7(3): 410-25.

[19] DjVu holds a great deal of promise in this area, however, as it is intended for raster image compression. File sizes are extremely small, and as the format becomes more pervasive DjVu may become a practical option for delivery of musical score images over the web.

[20] While many have hesitated to deliver PNG files over the web because of browsers' slowness to support the format, pre-version 4 browsers that do not support it are now miniscule at best, according to The Counter <http://www.thecounter.com/stats/2002/April/browser.php>.