INTERFACES FOR DOCUMENT REPRESENTATION IN DIGITAL MUSIC LIBRARIES

Andrew Hankinson[1] **Laurent Pugin**[2] **Ichiro Fujinaga**[1]
[1] Schulich School of Music of McGill University, Montréal (Quebec) Canada
[2] RISM Switzerland, Bern, Switzerland
{ahankinson,laurent,ich}@music.mcgill.ca

ABSTRACT

Musical documents, that is, documents whose primary content is printed music, introduce interesting design challenges for presentation in an online environment. Considerations for the unique properties of printed msic, as well as users' expected levels of comfort with these materials, present opportunities for developing a viewer specifically tailored to displaying musical documents. This paper outlines five design considerations for a music document viewer, drawing examples from existing digital music libraries. We then present our work towards incorporating these considerations in a new digital music library system currently under development.

1. INTRODUCTION

In 2008, the Swiss working group for the Répertoire International des Sources Musicales (RISM) project began work towards digitizing its national music collection. These digitized scores would be incorporated into an online catalogue of works, and would allow users of this system to view these items online. One crucial element for the success of this project was the implementation of software that presented these documents using musically consistent techniques.

The presentation of printed musical materials in an online environment poses interesting design challenges. Music and text documents have superficial similarities—they are written or printed on paper and bound in books—but they also differ significantly in their intended use, complexity of notation and stylistic considerations for presentation on the page.

When displaying music documents in an online environment these differences should be taken into account. Beyond putting the scanned content online, there needs to be consideration for how to show the material to users. As we will demonstrate in our literature review, the presentation of content can have a significant impact on a user's ability to navigate and comprehend the content itself. Next, we will propose five design considerations, formulated as requirements for implementation in a document viewer for a digital music library. Accompanying these design considerations we will show specific examples of how these have been implemented in existing digital library systems. Finally, we conclude with a brief discussion of our implementation of these design considerations, as well as commentary on possible directions for future work.

2. BACKGROUND

In his 1984 dissertation, Byrd [1] describes printed music—specifically, conventional music notation (CMN) —as a "modified coordinate system" that encapsulates semantic, syntactic, and graphic complexity occurring in four dimensions (pitch, time, loudness, and timbre). Accompanying the complexity of the music itself are practical considerations that play an integral role in the interpretation of the materials, such as page layouts, line justifications, and convenient page-turns. While these considerations are not part of the musical content, they are part of the total information content of the score. Put another way, while there is no one correct way to present the printed music, there are many wrong ways to present it that can lead to mis-interpretation of the music itself. The dimensionality and complexity of printed music, Byrd states, exceeds the complexity of printed text and is central to understanding the problems that exist with computerized analysis of these materials.

The delivery of information in online environments is an area of research that has received quite a bit of attention. In particular, Thong et al. [2] show that "[in] the context of digital libraries, it not only matters what we put on the screen, but how." They continue: "[the] way that information is arranged on the screen can influence the users' interaction with digital libraries beyond the effect of the information content."

Additional research has conclusively identified the affective relationship between the aesthetic perception of the materials and its effect on cognition and learning.

Kurosu and Kashimura [3] found that "[users] may be strongly affected by the aesthetic aspect of the interface even when they try to evaluate the interface in its functional aspects." Building on these findings, Tractinsky et al. [4] performed a study of automated teller machines and found "strong correlation between users' perception of an interface aesthetics and their perception of the usability of the entire system." They postulated that factors of aesthetics and usability can play a significant role in the overall satisfaction derived from an interface.

These studies' results are congruent with other work in the affective nature of human-computer interaction. When suggesting that "attractive things work better," Norman [5] (building on studies by Ashby et al. [6]) suggests that aesthetic interfaces can lead to a greater overall satisfaction in an interaction, which in turn can have significant effects on understanding the content. Increases in tension or anxiety, caused by unpleasant experiences with a system can negatively affect cognition of the material, leading not just to an unpleasant interaction, but also a decrease in the users' ability to understand the material itself.

While most usability research for digital libraries has focused specifically on textual materials, there has been work done on the evaluation of digital music library interfaces. Byrd and Crawford [7] touch on the topic of user interfaces for music information retrieval, simply stating that they are "hard." Byrd and Isaacson [8] address problems of music representation in a digital music library; however, they deal specifically with issues of notation layout, and not with interactions with digitized print materials.

The VARIATIONS project at Indiana University has conducted a number of usability studies on their system. Fuhman et al. [9] observed that non-musically trained users of their system took longer to complete musically-oriented tasks than musically-trained users, and gave a lower overall subjective rating to interfaces designed for displaying musical content. One possible explanation for this might be that musically-trained users have learned specific techniques for interacting with musical materials that are not shared by users who are unfamiliar with this content.

Finally, the SyncPlayer software [10] has received quite a bit of attention as a system that provides an easy-to-use interface for navigating score and audio representations of music. While this software presents an interesting interface for viewing and navigating complex scores, it was not included because it has not been used in a large-scale, public digital music library implementation.

3. DESIGN CONSIDERATIONS

As part of the design process for the document viewer, we identified five key considerations for designing an interface specifically for displaying printed music. These were formulated to encapsulate both the musical considerations of the documents, as well as some behavioural considerations of our target audience—musicologists and music researchers. For each consideration, we examined a number of existing systems used for displaying digital documents, musical or otherwise. By looking at these systems we were able to understand the current state of the art for displaying musical items, as well as discover interesting techniques to incorporate into our own implementation.

For a list of all the systems mentioned here, please refer to Table 1.

American Memory Project-Sheet: Music from the Civil War Era http://memory.loc.gov/ammem/cwmhtml/cwmhome.html
British Library "Turning the Pages" Project http://portico.bl.uk/onlinegallery/ttp/ttpbooks.html
Chopin Early Editions http://chopin.lib.uchicago.edu
Digital Image Archive of Medieval Music http://www.diamm.ac.uk
Google Books http://books.google.com
Inventions of Note Sheet Music Collection http://libraries.mit.edu/music/sheetmusic
Juilliard Manuscript Collection http://www.juilliardmanuscriptcollection.org
Lester S. Levy Sheet Music Collection http://levysheetmusic.mse.jhu.edu
Neue Mozart Ausgabe http://dme.mozarteum.at
Schubert Manusckripte http://www.univie.ac.at/wwtf/schubert
Sibley Music Library http://urresearch.rochester.edu/handle/1802/291
VARIATIONS Score Prototype http://www.dlib.indiana.edu/variations/scores
World Digital Library http://wdl.org

Table 1. Digital Music Libraries Examined

3.1 Preserve Document Integrity

One of the most common methods for presenting pages in a digital library is as a series of images on separate web pages, with navigation elements such as 'next' and 'previous' links, drop-down menus or hyperlinked page numbers as the primary means of navigating through the item. This method of document display suggests an 'image gallery' metaphor, rather than representing the

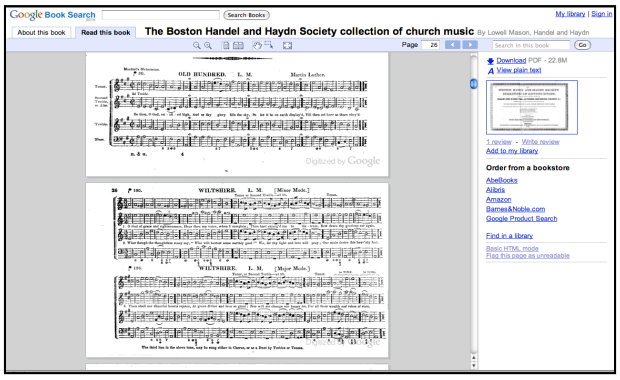

Figure 1. Google Books interface. The inner frame scrolls, with item metadata presented in the sidebar.

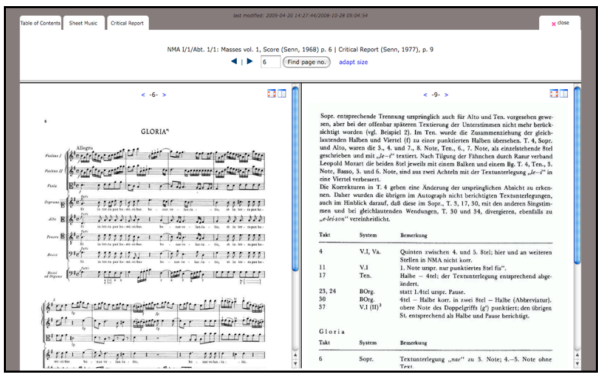

Figure 2. Side-by-side presentation of items in NMA. Musical material is presented on the left, and a critical report is on the right.

cohesive original document as a single entity. To preserve this cohesiveness, one of our design goals was to implement a display metaphor that preserved the original document integrity. Google Books, the Neue Mozart Ausgabe (NMA), and the VARIATIONS prototype viewer provide interesting examples of this functionality. These systems present the items as a single, scrollable entity embedded within a frame on the web page. This allows users to scroll very quickly through the item without having to click 'next' and 'back' links and wait for the page to reload.

A different technique was employed by the University of Illinois collection and the British Library's "Turning the Pages" project. These presented their documents using a book metaphor where users could use the mouse to 'turn' the pages. As a navigation system this was largely a novelty and presented some usability challenges for turning one page or many pages simultaneously. However, these systems excelled at presenting an accurate picture of the original page and book layout, an especially important consideration for musical materials.

3.2 Allow Side-by-side Comparison of Items

Musical documents can be divided into multiple physical items, with each item containing a portion of the complete musical work. Choir part books and orchestral instrument parts are common examples, but this can also extend to opera scores and libretti, early and later editions of a work, theory treatises and criticisms, adaptations, reductions, or various other modifications. It is not uncommon for scholars to need to consult multiple volumes for a single score.

Two systems, the Digital Image Archive of Medieval Music (DIAMM) and the NMA, had the facility for displaying multiple items, but neither of them allowed multiple musical items to be displayed simultaneously. DIAMM displayed corresponding scans from entries in a printed RISM catalogue that was digitized, while the NMA displayed scans from a published critical report on that piece of music (see Figure 2).

3.3 Provide Multiple Page Resolutions

When studying older manuscripts or printed works, the ability to view small details on a page, such as faint pencil markings or smudged note heads, can provide valuable information to the scholar. High-resolution images provides users with the ability to 'zoom in' on these markings, while lower resolution ones would allow them to move quickly through an entire document without having to navigate large pages.

Most of the systems examined provide more than one size of image. Typically, they would provide three page image sizes: a 'thumbnail' view for quick selection and browsing, a 'browser-safe' view for fitting in a browser, and a 'high resolution' view for downloading, printing, or further detailed study.

In two cases, DIAMM and the World Digital Library (WDL) provided methods for smoothly zooming in and out from the page images. In the case of DIAMM, they used an Adobe Flash-based viewer called "Zoomify," typically used for viewing high-resolution landscape photographs, while the WDL used a technology developed by Microsoft for their Photosynth viewer (see Figure 3).

3.4. Optimize Page Loading

Showing multiple high-resolution document pages presents significant challenges for network and browsing speeds. Furthermore, we know that our target user base often works in environments such as small libraries, monasteries, churches, or in rural locations, where bandwidth can be at a premium. To address these issues, one of our design goals was to only display the pages and areas of the page that the user was currently viewing. This would preclude the need to download an entire document of high-resolution page images if they

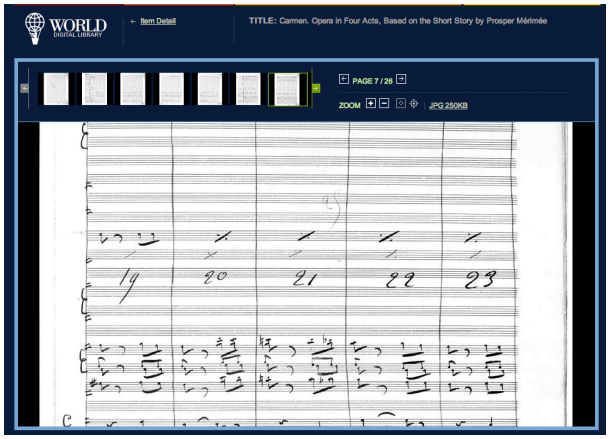

Figure 3. Zooming in on a manuscript in the World Digital Library

only wished to consult a single page.

As mentioned previously, Google Books uses an optimization technique that loads page images on demand. The Illinois Flip Book (beta) system seem to offer this as well, but at high zoom levels it required the user to download the whole high-resolution page image, slowing down the interaction.

The viewing system for the Schubert Manuskripte library used a segmentation system to display single high-resolution scans of a single page. Each page image was broken into smaller image tiles that could be downloaded in parallel, theoretically speeding up the interaction. However, it seemed to use real-time image manipulation (e.g., re-sizing and rotating) on the server side, meaning that any speed optimization gained in parallel download was lost while the user waited for the server to recalculate the image.

3.5. Present Item and Metadata Simultaneously

The catalogue record of a document often contains more information than is immediately available in the item itself or can serve to correct erroneous or outdated information on the item. For example, some compositions have been commonly attributed to the wrong composer, or their catalogue of works may have updated numbers. Although this seems like a small interface consideration, many implementations we examined would open images in the current or new window, replacing or obscuring the metadata and causing users to constantly flip between two browser windows or use the 'back' and 'forward' browser buttons to switch between item and item record.

The reasons for this separation are varied. Some systems, e.g., Harvard, DIAMM, the University of Illinois, and Juilliard, used document presentation software separate from its catalogue to display the actual item. Other systems, such as the American Memory Project and the Levy Sheet Music Collection, separated the catalogue records and the navigation of the pages in the item on different web pages. Still others, such as the Sibley Music Library and the Inventions of Note collection, simply provided their items as PDFs to download.

The Chopin collection offered a "tab" for switching between the score and the bibliographic interface, (Figures 4 and 5) but switching between the two did not maintain the users' position in the score, reverting them to the view of the title page. Google Books and the VARIATIONS prototype feature a sidebar with some cataloguing information present, but the full catalogue entry was on a separate page.

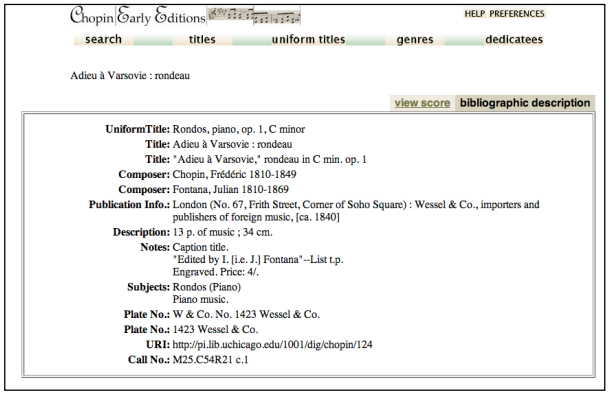

Figure 4. Bibliographic Description Tab in the Chopin Early Editions.

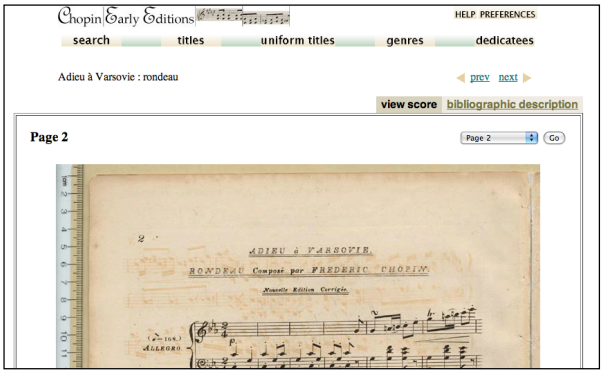

Figure 5. View Score tab in the Chopin Early Editions.

4. CURRENT WORK

While each design consideration we studied in our research can be found in several systems that we evaluated, our goal was to provide a system that would implement all of them. This viewer uses a number of technologies adopted from our examination of the existing solutions. Figure 6 shows a screenshot of our document viewer.

The unified document display methods found in the Google Books and VARIATIONS systems has been adopted. It has been enhanced to allow users to scroll both vertically and horizontally through an item, based

on the page orientation of the item (vertical scrolling for items in portrait orientation, horizontal for items in landscape orientation).

For musical works with many physical items, panels in the viewer allow users to view one to four items simultaneously. These panels are synchronized so that the location in a score is maintained across all panels as the user scrolls through one panel. The synchronization is currently limited to movement or section indexing that has to be provided by hand by the cataloger.

In order to provide multiple page resolution while keeping page loading optimized at high and very high resolutions (600 dpi or higher), the system uses a tiling mechanism that separate the images into small tiles, enabling it to serve only the displayed part of the document. By restricting the download to only the tiles that are needed by the user, we avoid the need to download the entire high-resolution image to view only a specific portion of an image. When combined with the unified document approach, this means that users can very quickly scroll through a document and zoom in on a specific page or set of pages without having to download the entire item.

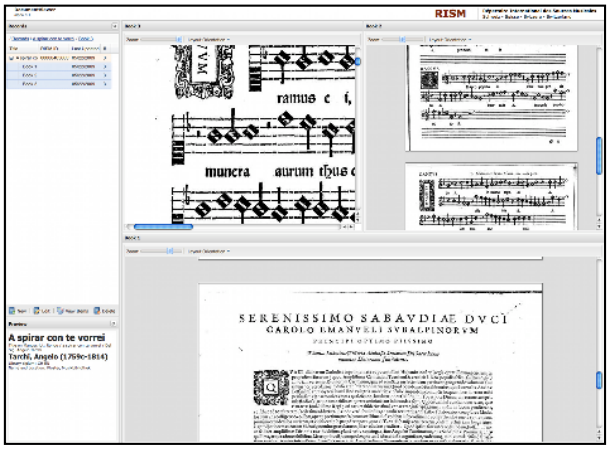

Figure 6. The Swiss RISM digital music document viewer. Three separate documents are displayed in panels in the middle of the page and can be scrolled vertically or horizontally. Document metadata appears in the lower-left panel.

Finally, the simultaneous presentation of metadata was incorporated into the interface by employing a sidebar similar to the VARIATIONS prototype. This panel can be hidden and shown dynamically, allowing users to concentrate on viewing the item but giving them easy access to the full catalogue record without having to navigate to another page.

4.1. Technical aspects

Ruby-on-Rails and MySQL provide the data storage on the server side. The client interface uses the ExtJS Javascript Framework [11].

For multiple page resolutions and optimized page loading, the system uses the IIP Image Server [12] tiling system. The image server separates large, high-resolution images into separate 256×256 pixel tiles and serves them on-demand.

Javascript Object Notation (JSON) is used as a communication language between the database, tile server, and user interface. Client and server communication is performed asynchronously. From a user's perspective, this means that there are very few page refreshes and performance approaches that of a native application instead a website.

To create new documents in this system, the images representing the page images are placed in a ZIP file and uploaded through the interface. These are then unzipped on the server side and processed using the VIPS image processing software [13] to create a pyramid TIFF [14] file that contains a sequence of images at increasingly coarse resolutions, representing zoom levels for these images (see Figure 7).

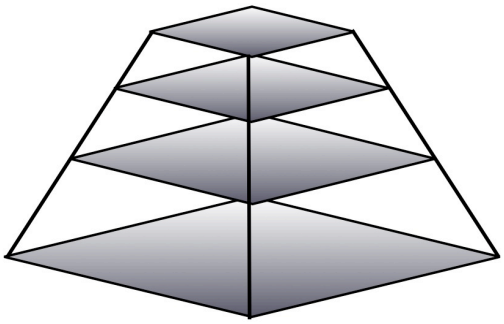

Figure 7. Pyramid TIFFs contain multiple resolutions of an image in a single file.

When a user requests a document through their web browser, the interface translates this into a request for the images and zoom level of the pages currently visible in the viewer. The images are then served to the client as separate tiles and re-assembled as a page image in the interface. This is repeated for each page so users can scroll through an entire document at very high resolutions without waiting for the whole document to download.

The software and display techniques presented here will be incorporated into a modular digital library system currently under development. Each component of this system uses open-source software, and we will be releasing this system under an open-source MIT license, freely available for implementation in existing digital library systems.

5. FUTURE WORK

The design considerations proposed here have not been verified by user testing. As part of our ongoing work on this software, we plan to study the impact of these design considerations on our target audience in real-world usage situations.

As part of the ongoing Swiss RISM project, the viewer interface will be integrated into the catalogue of Swiss musical works available online [15] as the documents are digitized.

6. CONCLUSION

This paper introduces our work towards a viewer for digital music documents, taking into account the unique properties of printed music and the expectations of users who use these systems. We also expect that having an architecture designed specifically for music documents will be of great benefit in the long run as it should facilitate the integration of other information research technologies specific to music, such as content-based synchronization or online optical music recognition.

7. ACKNOWLEDGEMENTS

This work was funded by the Swiss National Science Foundation and the Swiss Academy of Humanities and Social Sciences.

REFERENCES

[1] Byrd, D. 1984. Music notation by computer. PhD diss., Indiana University.

[2] Thong, J., W. Hong, and K.Y. Tam. 2002. Understanding user acceptance of digital libraries: What are the roles of interface characteristics, organizational context, and individual differences? *International Journal of Human-Computer Studies* 57 (3): 215–42.

[3] Kurosu, M., and K. Kashimura. 1995. Apparent usability vs. inherent usability: Experimental analysis on the determinants of the apparent usability. In *Proceedings of the Conference on Human Factors in Computing Systems,* 292–93.

[4] Tractinsky, N., A. Katz, and D. Ikar. 2002. What is beautiful is usable. *Interacting with Computers* 13 (2): 127–45.

[5] Norman, D. 2002. Emotion & design: Attractive things work better. *Interactions* 9 (4): 36–42.

[6] Ashby, F., A. Isen, and U. Turken. 1999. A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review* 106 (3): 529–50.

[7] Byrd, D., and T. Crawford. 2002. Problems of music information retrieval in the real world. *Information Processing and Management* 38 (2): 249–72.

[8] Byrd, D., and E. Isaacson. 2003. Music representation in a digital music library. In *Proceedings of the Joint Conference on Digital Libraries,* 234–36.

[9] Fuhrman, M., D. Gauthier, and A. Dillon. 2001. *Usability test of Variations and DML prototypes.* http://variations2.indiana.edu/pdf/VariationsTest.pdf [Accessed 20 May 2009].

[10] Fremerey, C. 2006. SyncPlayer: A framework for content-based music navigation. Diplomarbeit, Univ. of Bonn.

[11] ExtJS Framework. http://extjs.com/ [Accessed 9 August 2009].

[12] Pillay, R., and D. Pitzalis. 2008. *IIP Image Server*. http://iipimage.sourceforge.net [Accessed 20 May 2009].

[13] Cupitt, J., and K. Martinez. 1996. VIPS: An image processing system for large images. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE)* 2663 (19): 19–28.

[14] Library of Congress. 2009. TIFF, Pyramid. http://www.digitalpreservation.gov/formats/fdd/fdd000237.shtml [Accessed 22 May 2009].

[15] RISM Switzerland Online Catalogue. http://www.rism-ch.org/ [Accessed 9 August 2009].