Integrating Paper and Digital Music Information Systems

Karen Lin and Tim Bell

University of Canterbury, Christchurch, New Zealand

Abstract

Active musicians generally rely on extensive personal paper-based music information retrieval systems containing scores, parts, compositions, and arrangements of published and hand-written music. Many have a bias against using computers to store, edit and retrieve music, and prefer to work in the paper domain rather than using digital documents, despite the flexibility and powerful retrieval opportunities available. In this paper we propose a model of operation that blurs the boundaries between the paper and digital domains, offering musicians the best of both worlds. A survey of musicians identifies the problems and potential of working with digital tools, and we propose a system using colour printing and scanning technology that simplifies the process of moving music documents between the two domains.

Keywords: user interfaces, user needs, optical music recognition

1. Introduction

Traditionally musicians have stored and retrieved music scores using paper-based systems. Many musicians have built up personal libraries of music books, compositions, arrangements and sheet music. The acquisition or creation of documents is straightforward, but the retrieval or modification of scores is hindered by the inflexibility of the paper medium. A digital music library would have a number of benefits, including convenient retrieval (instead of searching through piles of music), ease of processing (such as part extraction), and communication (sending electronic copies to other performers). However working with digital documents also poses significant barriers for users more familiar with traditional paper documents. In this paper we explore the relationship between the paper and digital domains, and the possibility of allowing easy conversion between the two to allow documents to exist in both domains and be processed in whichever domain is the most convenient.

We begin by proposing a three-state model that provides a framework for reasoning about the different domains. Then we report on a survey of musicians to determine the kinds of retrieval and processing tasks they carry out, and their preferences for carrying out these tasks in each domain. We then report on a system in which color manuscript is used to simplify the transition between paper and digital media.

2. The states of a music document

Rather than just use two states of a document (paper and digital), we propose a model that divides the digital state into two: image data and semantic data. Figure 1 shows the resulting three-state model, with transitions showing how data can be converted from one state to another. This model is intended to help us to reason about the representation of musical scores. Digital image data is typically a bitmap representation of the pixels, whereas the digital semantic data stores information about pitch and rhythm. This acknowledges the difference between, for example, a scanned image and one that has been recognised so that meaning is associated with the symbols. Methods for converting between the three states are also shown. For example, digital image data can be converted to digital semantic data using an OMR (optical music recognition) system.

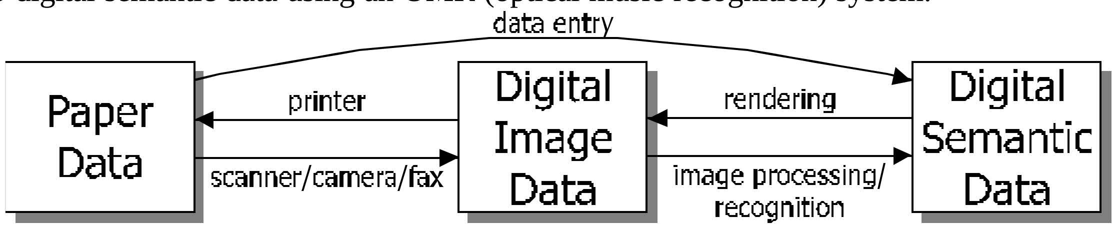

Figure 1: Three-state model of data

A document may well exist in all three formats at once. For example, a score in a musician's library might be scanned as a digital image and stored on a portable device, but also be recognised as best as possible to allow semantic information to be stored for convenient retrieval. This kind of approach is already common with textual documents. For example, the Adobe Acrobat Capture system [Merz 1997] can be used to put a paper document on-line in a form displaying the original scanned image, but 'hiding' an OCRed version behind it to allow searching and copying of the text. The more accurate the recognition, the more useful the retrieval process will be. Retaining the image data prevents mistakes in recognition being reflected in the displayed version of the document. Paper documents are still widely used for writing, storing, practising and performing music. Paper is reliable, inexpensive, has a high resolution, is easily marked up, and is very portable. We are interested in removing the boundaries between the three states so that, for example, an arranger could write a score on paper, scan it, print out an edited version, annotate it with pencil, and scan it again to have the differences incorporated in the digital version of the document.

3. How musicians manage music documents

After some preliminary interviews with professional musicians we designed a survey in order to assess how digital techniques could aid musicians with music information management. The musicians interviewed and surveyed came from a variety of backgrounds from classical to contemporary music. We were interested in musical activities of all kinds including composition, arrangement, performance, teaching, musicology, recording, accompanying, transcribing and so on. We asked about the tools that they use, their working procedures, difficulties that they face, and ways that digital storage and retrieval might be applied.

The survey was distributed to about fifteen Internet newsgroups. There were 40 respondents from differing backgrounds. Twelve of the 40 were full-time musicians, and 75% of the respondents perform musical tasks for at least ten hours per week. All the respondents peform multiple tasks; for instance, some are composers, arrangers and performers. It was interesting to find that six respondents compose exclusively by hand and seventeen respondents use both paper and computer software. This means that over half of the respondents still use paper for composition and arrangement, despite the survey being biased towards computer users because it was conducted on the Internet.

The survey identified a number of existing problems that with the use of computers for managing musical documents. These are as follows:

The cost of computer hardware and software is one of the main concerns for at least 30% of participants. Expensive hardware such as fast computers for running modern software, high-quality printers for a good printing result, CD-writers and hard disks for storage, as well as synthesizers and expensive music notation programs are beyond the means of many musicians. Moreover, rapidly changing software and hardware make the problem worse.

The size of a computer screen is another problem. Even though 17-inch screens are relatively common and larger screens are available, this is still very limiting for an orchestral score that includes several parts for different instruments. Some participants mentioned that it would be very helpful to have an affordable flat screen as big as an ordinary desk. Another problem is that screens can be tiring to look at for long periods.

Using computers can cause a loss of inspiration and creativity. This was an issue mentioned twice in the interviews, and by five respondents from the Internet survey. The computer imposes an extra cognitive step between having an idea and recording it in a document. Some musicians overcome this with experience, but at the cost of lost productivity while becoming accustomed to the new method. Furthermore, with handwriting any symbol can be placed anywhere on the page, whereas an infrequently used symbol (such as a glissando) may be very difficult to invoke in music processing software, leading to loss-of-activation errors for the user.

Loss of efficiency is experienced by some computer users because it can take more time to enter music into music software. Some musicians said that they could write music by hand several times faster than using a computer; one of them claiming it could be ten times faster. There is some research supporting this statement: Anstice [1996] found that hand-writing music takes about one-third of the time. Improving software by employing flexible input techniques may reduce this problem, and an experienced user might become quite efficient. One of the participants mentioned that he could enter music on the computer very quickly — he has used that software since 1985.

Even the start-up time for a computer can be a problem. Several participants said that when they had an idea, they would be unlikely to be bothered to turn on a computer and run music notation programs to compose the music. By the time they have waited several minutes for the computer and software to start, they have forgotten the music they thought of and the inspiration has flown. Other musicians said that although they do not write music by hand often, they particularly prefer to write this way when they are in a rush or some excellent ideas come to mind suddenly. They agree that grabbing a piece of paper to write music down by hand is the quickest way under such situations.

The long learning curve required by complicated and badly designed software was mentioned by several participants. Musicians have to spend their valuable time learning software, rather than creating music.

Software compatibility was another issue raised. One participant found that one new version of software had a different data format to the old version. Consequently, he had to enter all of his music again. Other respondents had similar experiences. One said that he has kept an ancient computer just to be able to print out his master's thesis composition. Compatibility is also a problem between different vendors, particularly in the absence of widely used standards for notation representation.

Hand-writing music on paper first also poses problems. As mentioned above, half of the participants write music by hand, and some then further develop or arrange their idea on the computer. In order to do so, they must enter the already-written music by using computer or synthesiser keyboards. This creates double handling. In particular, when the composition is large, such as an orchestral score, the whole process of moving handwritten music from paper to computer becomes very cumbersome. One musician with two children finds quality time is precious and limited. Once she completes an arrangement or composition, she does not want to take the time to copy or format the parts. She just wishes someone or something could take it from there and finish the routine work.

We have also observed that another problem with paper documents is the choice of writing implement. A pen is better if the music is to be photocopied and distributed, but a pencil is more flexible for correction of errors. Computer manuscripts offer the best of both worlds: they photocopy well, yet they are amenable to changes including copying phrases, transposition, and part splitting. Moreover, music written on the computer can be played back for 'proof reading' for rehearsal.

We were also interested in which criteria are most useful for searching for a piece of music in personal collections, music libraries or music shops. In the future the distinction between these is likely to be blurred by digital libraries. The survey showed that composer, title and instrument would be the most popular criteria for searching a digital music library. These can all be extracted from the documents' text, without using Optical Music Recognition.

One musician said that he would like a database collection that categorises music according to musical content – time signature, type of piece, source, and melodic characteristics. When he gets results from searches he would like to be able to see the themes, not just their titles. Another musician would like a completely customizable database. Other ideas suggested for a music database are: recording files in a database directly linked to the midi software, comparing audio input and scores to find a title, and entering a melody on a keyboard and have the computer search for close matches. We note that some of these features are already available in experimental systems such as the New Zealand Digital Library [Bainbridge et. al. 1999].

We also asked the musicians to suggest improvements to existing music information systems.

One musician surveyed pointed out the difficulty of scanning bound sheet music and books. A fast system for automatically scanning these would be invaluable for converting from the paper domain to digital form. This might be achieved by a high-resolution overhead camera that could take 'snapshots' as the user turns the pages.

Many of the musicians surveyed would like a page turning device. On printed music, page breaks occur during a rest or one-handed section if possible. However, if there is no suitable position, unless playing from memory, musicians are required to stop playing to turn the page. In an orchestra the music is typically shared between two players, so one stops to turn the page, resulting in a 'dip' in the music level [Graefe et. al. 1996]. It would be very helpful if there were a digital music display with a system that could follow the music and do the page turning for the performer. Some research has been conducted [Graefe et. al. 1996, McPherson 1999, Bellini 1999] but the quality, bulkiness, cost and reliability of computer screens make a paper based system more attractive at present.

Another problem raised was the unreliability of scanning and recognising hand-written scores and old imprints. Even a system that is accurate 99% of the time could have an error in every few bars of a complex piano piece. The time taken to find and correct such errors may not be much less than the time to enter the music correctly by hand in the first place [Hewlett and

Selfridge-Field 1994]. For older music or handwritten music being studied by a musicologist, existing OMR systems are not able to do a very accurate job.

Other things that the musicians surveyed would find useful include: a virtual accompanist; video and music being integrated more elegantly; compression technologies improved to save hard disk space; more efficient and stable software; and a 'brilliant' pen input system.

4. Music processing using colour

From the surveys it was clear that composers and arrangers would benefit greatly from a system that allows them to write music by hand and later transfer it to the digital domain. The recognition of handwritten music is difficult, but if the writer knows that the music will be transferred to a computer at a later stage, then the manuscript and notation can be adapted to increase the chances of this being done accurately.

Previously we had explored the use of a pen-based computer so that the computer was able to track all writing [Anstice et. al. 1996, Ng et. al. 1998], but this required specialist equipment that to date has not become commonplace.

Another approach is to use color on the manuscript to simplify the task of separating symbols from the stave line [Lin and Bell 1999]. A major source of error in OMR systems is trying to determine the shapes of objects superimposed on a stave. A simple and effective way of using color is to print the manuscript lines in a color (such as red), so that color as well as position can be used to distinguish any handwritten symbols. Since inexpensive color inkjet printers and color scanners are available and widely used, this method is readily available to any user who has access to a fairly conventional computer setup.

However, identifying colored staves is not as simple as matching pixels of a given color, as the color of a scanned line can be quite variable, and imperfections in printing the image can mean that a line that appears to be red can contain many different colored pixels. Likewise, grey or black writing on the stave can vary in color, and where it intersects with the colored lines there can be smudging (where the colored ink is picked up by the pen or pencil) or blurring (where the ink shows through the superimposed symbol).

To avoid these problems we use traditional stave line identification techniques to provide clues about where the boundaries of symbols are expected to be. The variation in colors is overcome by mapping the scanned colors into a color space that distinguishes the key features of the different colors we are looking for.

Figure 2: Two sample pieces of music with blue staff lines

Figure 2 shows two sample pieces of music that have bright blue staff lines and grey notation written using a 4B pencil. The bright blue staff lines were printed one point thick with the bright blue color (RGB 0,0,1), and then printed out onto white paper by an inkjet printer at a resolution of 300dpi. The notation was written by hand using a 4B pencil, and then the image was scanned through an inexpensive color scanner with default scan settings at a resolution of 300dpi.

There are three colors that need to be distinguished in the image: the white background, the grey symbols, and the blue lines. In the scanned image there is substantial variation in the scanned colors of pixels belonging to all three of these sets.

We found it useful to transform the colors to the HSV (Hue, Saturation, Value) color space, which matches human perception better than the RGB (Red, Green, Blue) space used by scanners and screens, or CMYK (Cyan, Magenta, Yellow, Black) used by printers. The Hue identifies a color of the spectrum, and the Saturation indicates how intense the color is. The Saturation is particularly useful for distinguishing the blue stave lines because a value significantly greater than zero indicates that a pixel is 'colorful' instead of being a shade of grey (from black to white). A nice side effect of using saturation is that it does not matter which color the stave lines are as long as they are not a shade of grey.

Experiments have shown that a saturation of more than 20% indicates a colored stave, while 2 to 10% indicates pencil writing, and less than 2% indicates the white background.

We also found it useful to convert an image to the CMYK color space to get additional information. The K (Black) value indicates the 'darkness' of a pixel, and was useful for identifying the writing that tended to have a higher K value. Because CMYK uses four parameters to describe a three-dimensional space, it is under-constrained, and several different formulae are available to calculate K. The most trivial is to use K=0, since the three other colors, CMY, also cover a valid color space, although this provides no help for our system. The maximum value that could be used for K is the minimum value of C, M and Y in a CMY representation. This value is useful, although we found that some of the more sophisticated formulae used by commercial systems worked even better. These formulae are not widely available because they are commercially sensitive, and could only be inferred from samples of colors given to the systems.

Figure 3 shows the result of using these techniques to extract the pencil symbols from the blue stave-lines of Figure 2. We can see that the shapes of the grey notations have been almost entirely preserved and the stave-lines are almost completely removed. The accuracy is even better if a black pen is used instead of pencil, although this is less flexible for the user.

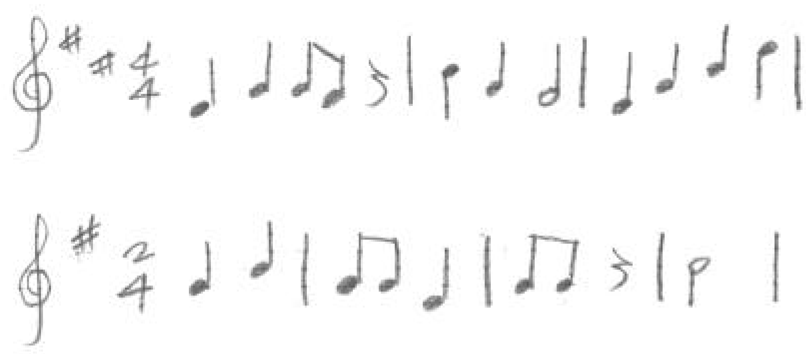

Figure 3: The sample of music in Figure 2 after the stave-lines have been removed

We have also experimented with using color for the annotation of music, and have found that similar processing techniques to those described above can be very effective for extracting notation. For example, and arranger might use a colored pen to mark a section of music to be copied, deleted, or transposed.

5. Conclusion

There is a demand for merging the paper and digital domains for musicians who would like to be able to take advantage of the ease of use of paper documents, and the flexibility and retrieval properties of digital documents. Giving composers and arrangers the option of working with pencil in the paper domain eliminates the cognitive hurdle of using software while engaged in a creative endeavour. The survey has identified a number of features that would be useful in systems that operate in both domains. By using colored stave-lines the conversion to digital semantic data is provided with extra information to help with the accurate removal of stave-lines, and to help identify special symbols.

References

J Anstice, T Bell, A Cockburn and M Setchell: The design of a pen-based musical input system, OZCHI 96 *(Hamilton, New Zealand)* (1996) 260-267

D Bainbridge, C Nevill-Manning, I Witten, L Smith and R McNab: Towards a digital library of popular music, The 4th ACM conference on Digital Libraries, (*Berkeley)* (1999) 161-169.

P Bellini, F Fioravanti and P Nesi:: Managing Music in Orchestras, IEEE Computer, 32(9), (1999) 26-34

C. Graefe, D Wahila, J Maguire and O Dasna: Designing the muse: A digital music stand for the symphony musician, CHI 96 *(Vancouver, BC, Canada)* (1996) 436-441

W B Hewlett and E Selfridge-Field: How practical is music recognition as an input method?, in *Computing and Musicology: An international directory of applications*, CCARH, Stanford, California (1994) 159-166

K Lin and T Bell: Music processing using colour, IVCNZ99 *(Christchurch, New Zealand)* (1999)

J McPherson 1999: Page Turning — score automation for musicians, Honours project report, Department of Computer Science, University of Canterbury, Christchurch, NZ (1999)

T Merz: Postscript and Acrobat PDF, Springer-Verlag, Berlin (1997)

E Ng, T Bell and A Cockburn: Improvements to a pen-based musical input system, OZCHI 98 (1998)