

VISUAL COLLAGING OF MUSIC IN A DIGITAL LIBRARY

David Bainbridge

Dept. of Computer Science
University of Waikato
Hamilton
New Zealand

Sally Jo Cunningham

Dept. of Computer Science
University of Waikato
Hamilton
New Zealand

J. Stephen Downie

Graduate School of Library
and Information Science
University of Illinois
Urbana-Champaign, IL, USA.

ABSTRACT

This article explores the role visual browsing can play within a digital music library. The context to the work is provided through a review of related techniques drawn from the fields of digital libraries and human computer interaction. Implemented within the open source digital library toolkit Greenstone, a prototype system is described that combines images located through textual metadata with a visualisation technique known as collaging to provide a leisurely, undirected interaction with a music collection. Emphasis in the article is given to the augmentations of the basic technique to work in the musical domain.

1. INTRODUCTION

Looking at patterns of use in real music collections—physical libraries, music stores—we see people (of their own volition) spending a significant amount of time looking, browsing, sifting through the music/documents [5]. This is in stark contrast to observed behavior in digital libraries; transaction log analysis reveals brief sessions, few searches, short queries, and few documents viewed [7, 12]. Is this a true reflection of how users, in general, wish to interact with digital repositories or merely an artifact of browsing mechanisms that are ineffectual or too cumbersome to be used? This issue has led to various research initiatives developing richer forms of support for browsing with text-based digital libraries, for instance [3, 15, 14]. Given a strong desire shown by people to browse real music collections then the ability to support this activity within a digital music library seems especially important.

This provided the motivation for the work described in this paper: the adaptation of a visual technique known as collaging to the music domain. Collaging is a technique devised by Kerne [9] for use in the more general environment of the web. In one example of its use, a user keys in several web domains—say mtv.com, archive.org/audio, and ceolas.org—presses *go* and then observes a visualisation of images drawn from crawling these web sites con-

currently. Images already in the visualisation fade over time and as new images are retrieved from the various sites, they are added in to the visualisation on top of what is already there. If at any stage a displayed image sparks some interest, the user can click on it and have a new web browser open up displaying the web page the image comes from.

This paper is organized as follows: Section 2 reviews related work in collaging and browsing in digital libraries, and the “laid back” approach to interacting with information collections; Section 3 presents a walkthrough of a prototype music collection that includes collaging; Section 4 describes implementation issues for creating the prototype; and Section 5 is a summary of our work.

2. RELATED WORK

Previous graphics-based browsing tools for music document collections primarily support visualization of entire collections; for example, as a self-organising map [14] or a 3D vector space [15] showing documents clustered by genre. While these visualisations can give an immediate, powerful overview of relative size of groupings and a sense of distance between clusters, they are less effective in supporting browsing of the documents themselves, since the idea of an overview means that the documents must be minimised and represented at the top (browsing) level in a relatively impoverished fashion (for example, by a single point in a 3D space) [3].

Earlier work applying collaging to information seeking focused on using the resulting collage as a search tool, for example to build a collection of Web documents based on a CollageMachine search/browse over Google [8]. Based on user evaluation studies, collaging seems ill-suited to supporting more formal information seeking activities such as direct search; users objected to the lack of predictability in the collaging, and the lack of transparency in the relationship between the search terms that seed the collage and the collaged “results.” Collaging appears best suited to providing a browsing experience that is engaging and interesting, but that does not require the direction and focus on the part of the user that is characteristic of searching or even of other types of browsing (for example, browsing by metadata or by genre). This type of activity, where the system selects information for viewing

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.
© 2004 Universitat Pompeu Fabra.

based on little (or even no!) control from the user, has been termed “laid back,” to distinguish it from the more active and user-directed “sit forward” information seeking activities [6].

The utility of the laid back approach to browsing has supporting evidence in the HCI, cognitive science and psychology of memory literatures. It is now a well-known trope in these literatures that “recognition” is greater than “recall” [13]. Since laid back browsing is premised on the user recognising an association presented by the collaging interface rather than having to recall a specific fact, melody, or set of appropriate search terms, it follows that the collaging technique, and its laid back approach, should provide a beneficial browsing enhancement for music seeking users.

Previous implementations of collaging have been directed at using extracted images or text snippets to explore or search document collections that are primarily text [8, 10] or images [2]. How appropriate is this technique likely to be when applied to collections of audio documents? In answer, we note that the visual collage-based browsing we describe in this paper is surprisingly close to a common browsing activity observed in music stores—flicking through horizontal displays of music CDs. In this latter activity, shoppers glance at the CD cover images as they are flicked past, occasionally pausing to look more closely at an item of interest, perhaps to read its cover notes and to “hear the song in my head” [5].

The use of visual surrogates for audio files in a collage also seems appropriate given evidence that appearance plays a significant role in the organization and display of personal music collections—the attractiveness of the CD covers, the arrangement of CDs on shelves, and the general effect achieved by prominently positioning the collection as the centrepiece of a room [4]. Collaging is designed to support the display of documents in an interesting and attractive manner, with the added benefit that the juxtaposition of documents in a novel and serendipitous process may create compositions that the browser finds meaningful [11].

3. PROTOTYPE

For pedagogical purposes a Greenstone digital library collection was built based on a user’s (modest) personal collection of 80 MP3 files. The result can be seen in Figures 1 through 4. In addition to the commonplace browsing structures—such as songs arranged alphabetically by Title or grouped Artist, for instance—the music digital library can provide additional functionality and form. In a moment we will look at one of the more novel ideas, collaging tailored to music; here we will briefly mention some more prosaic features that come from the existing textual support within the digital library toolkit.

Figure 1 indeed shows the familiar “browse by artist” listing. Next to the bookshelf icon in the figure is displayed in brackets the number of songs by that artist, which if clicked upon opens up to show the list. Further clicking

on a particular item will result in the song being played.

The ability to search document content is one of the pillars upon which digital libraries are based, and although Greenstone does not currently include support for direct content searching of raw audio files,¹ in this context the metadata stored in ID3 tags makes for a useful foundation upon which to base textual queries. The digital library toolkit is flexible over what indexes can be formed: for this collection it was decided to index titles, artist, and a conglomeration of all text fields extracted from a file (akin to full-text search). Field searching is also supported through an advanced search page.

Depending on their source, not all MP3 files have their metadata fields filled out appropriately. Quality control is a hallmark of digital libraries that distinguishes them from the web at large. To help with selection and maintenance of documents in a digital library, Greenstone provides a graphical interface for adding and editing existing metadata, along with managing other workflow aspects of a digital library lifecycle [1]. Using the graphical tool a user can shape their music digital library to their taste. Metadata need not be restricted to the fields present in the MP3 files. A system of metadata sets is supported that allows the use of existing standards, such as Dublin Core, as well as ad hoc sets developed by a user or community of users. Through the augmentation of metadata fields, richer searching and browsing structures can be developed, for instance yielding a subject hierarchy for the collection.

Figure 2 shows a more novel form of browsing the collection. During the building phase of the digital library, title and artist information contained in each MP3 file has been used to locate two images deemed to be relevant to the musical work. Collectively, these form the basis of the collage. In broad brush strokes, one image at a time is randomly taken from the set and placed at a random location within the visualisation. As time passes, the image begins to fade until, eventually, it can no longer be seen, at which point it is removed from the visualisation and added back in to the set of possible choices. In parallel with this, new images are being added at regular intervals. The selection method of acquiring images is heuristic based, which we explain more fully in the next section. Album covers, photographs of the actual vinyl disc—some younger users of the digital library collection may need this format explained!—and publicity shots of artists are most common.

In Figure 2 the most prominent images in the collage are two record covers: The Beatles/1967–1970 (colloquially known as the blue album) and in front of that the single’s cover for *Another one bites the dust* by Queen; to the right of these there is also a picture of David Bowie live in concert, although this is not so clearly reproduced as a monochrome figure. If the user clicks on a particular image, then a new browser window opens up, as is shown in Figure 3 where the aforementioned David Bowie pic-

¹ Greenstone does however support direct content querying of symbolic music.

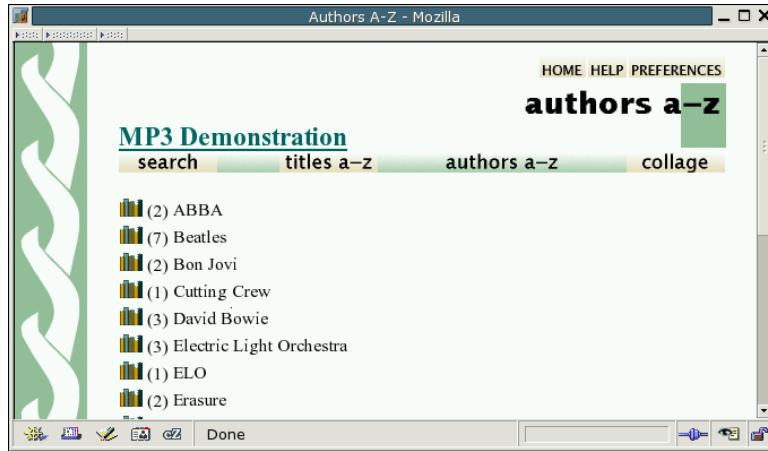


Figure 1. Browsing collection by artist



Figure 2. Collaging in action.

ture has been targeted. The new window is a metaphor for the selected record, and shows the two images associated with the track (fashioned on the idea of opening up the CD cover) while the audio is played through an embedded MP3 applet.

Additional resources relating to the musical track are hyperlinked through this page. If the user clicks on either of the images, they are taken to the web where the image was sourced from. In Figure 4 the user has clicked on the right-hand image of the David Bowie ‘record’ and brought up series of interviews with the artist. In the case of *Take a chance on me* by ABBA, to take another random example, the images take you to a discography site for the band. The title of the song, displayed below the images, is also hyperlinked and is used to automatically formulate a query on Google Images to find, with a certain degree of accuracy, related items. There is also a small embedded Java applet that plays the track, primed to start once the page loads and to stream the audio if possible. Follow-

ing that there is an MP3 logo that links to the original file. This allows a user to download the file if they wish or play it using a different application. This last detail in itself is not especially astounding, but does help emphasize that digital libraries need to integrate with other aspects of the user’s computer environment.

4. IMPLEMENTATION

For rapid prototyping, Greenstone—a versatile open source digital library toolkit that the authors have considerable experience with—was chosen as an implementation framework [16]. The toolkit incorporates a flexible architecture for digesting new document formats and customising how they are indexed, browsed and presented, making it ideal for our needs [17]. The architecture has four main components: 1) an internal XML document format to which all documents are converted when they are imported into the system (native media such as an MPEG video is bound



Figure 3. Viewing the work: two images, title and the audio playing

to the format through a linking mechanism); 2) a set of parsers called plugins that process document and metadata formats; 3) a set of modules called classifiers that are used to build browsing structures; and 4) a scheme for designing individual digital library collections by providing a configuration file that specifies what kind of documents and metadata they are to contain, and what searching and browsing facilities they provide.

To develop the music based prototype, a plugin for the MP3 format was needed; its primary purpose to bind the source MP3 file and digest the metadata stored in its ID3 tags into the XML representation. To support the collaging metaphor, a new classifier module was also required along with an applet to drive the visualisation. We describe each of these in turn.

4.1. An MP3 Plugin

Plugins in Greenstone are written in Perl, and are accessed when a digital library collection is built. A substantial array of plugins already exists, processing formats as varied as HTML, Word, PowerPoint, PDF, Images, Email messages, Excel spreadsheets and MARC data to name a few. There is also a plugin known as the UnknownPlug which allows a digital library designer to digest, in a rudimentary fashion, documents of a type that do not have a specifically targeted plugin by setting basic metadata such as the document's filename and file size, and by binding the native file into the XML format.

UnknownPlug was the starting point for writing the MP3 plugin, which—inheriting from this plugin—is essentially a tailored version of this, where more detailed metadata information, such as song title and artist, is extracted from the file. The lion's share of this work, in turn, is implemented by the MP3::Info package, available through the Comprehensive Perl Archive Network (www.cpan.org). Retrieving an associate array of values through this package, each entry is added into the Greenstone system as metadata bearing the same tag label, prefixed by 'mp3' as an XML namespace.

With an eye on supporting collaging at runtime within the digital library, an additional routine was added to the



Figure 4. Original web page from which second David Bowie image sourced.

plugin that used the title and artist information to perform a search on Google Images. To be precise, two exact phrases are formed based on the file's artist and title metadata and a search, restricted to JPEG files only, is issued to Google Images. If for some reason the file is missing either of these fields, an alternative search is generated based on the audio track's filename, which due to informal conventions that have grown up around raw audio files, often encodes similar title and artist information. Elementary manipulation of the filename is performed, such as converting underscores (_) to spaces which are often used as a substitute in filenames.

For the resulting images identified by the Google search, the first two valid matches are mirrored into the collection's document import area. Since the web is inherently a noisy medium, even if the matching image is available, it is sometimes the case that it is still not a valid JPEG image. To help spot such anachronisms, we use *identify*, a command-line utility included in ImageMagick's suite of open source image manipulation tools (see [imagemagick.org](http://www.imagemagick.org) for more details). The necessary Perl code to implement this totalled around 300 lines. Similar work could be done for comparable audio formats such as Ogg Vorbis and Windows Media Audio (WMA).

4.2. A Collaging Classifier

To develop the collaging facility within Greenstone a mixture of Perl and Java code was needed. The Perl code implemented a new classifier that linked document identifiers of musical works in the digital library to the associated images located through Google Images. This was not an unduly complex task as much of the required functionality resides in a base class for classifiers. Roughly 200 lines of code was needed to develop the necessary structure. Care was given to capping the number of document identifiers clustered together as one block, for runtime efficiency. If the blocks get too large, then there is a no-

ticeable delay detectable in the user interface between the information being retrieved and the resulting images being displayed.

The Java code took the form of an applet to provide, within the context of a web-browser based digital library, the necessary fine-grained interaction. The applet was multi-threaded, with one principle thread accessing the information stored by the Collage classifier: once furnished with a document's identifier, the applet can make subsequent calls to access metadata and associated files. Since the images associated with each MP3 file reside on the Greenstone server, these need to be downloaded locally by the applet for use in the visualisation. This too took place in a thread. Finally, a third thread was used to manage the graphic canvas in which the collage visualisation was presented, set to refresh at regular intervals.

Alpha values—a technique in graphics to control the transparency of pixels within an image—was used to implement the fading of images over time. Earlier we described the placement of images as random. This is an over-simplification. To make better use of the real estate available within the visualization, a data-structure is maintained recording the areas currently occupied by the images in the display and how faded they are. When selecting a new “random” position to place a fresh image, several random positions are calculated and the one that resides in the most dormant area is chosen.

Queues are kept of the downloaded and currently displayed images. Parameters to the applet control aspects such as the dimensions and background colour of the canvas, and for conservation of memory (if desired), how large the various queues can grow. Compared with the code written for the new plugin and classifier, the applet was a more substantial—but still modest in the larger scheme of things—undertaking requiring 2,500 lines of code.

4.3. Runtime

The final component for developing the prototype within Greenstone was to specify how “the works” (our MP3 files) were to be displayed. This is an aspect of digital library design that varies greatly depending on the types of documents contained in a collection. Control of this in Greenstone is exerted through format statements, a mechanism for customising—on a collection by collection basis—presentation and functionality.

For the display of MP3 target documents, it was decided to rekindle the notion of a person perusing the album cover sleeve while listening to the music, albeit in a simplistic form. We have already seen the result of this in Figure 3. It was accomplished by writing a format statement that sets out a table in HTML with the two sample images (which, incidentally, are stored as metadata for the document by the MP3 plugin at build-time) in the first row, and the artist, title (more metadata) and embedded MP3 Java applet in the second.

Additional syntax was included to hyperlink images and title information to external resources on the web. The intention here is to engage the listener in related material.

An anecdotal example of this in use was with one user who, on test-driving the prototype, remarked they'd always wondered if a particular song by The Beatles was *I want to hold your hand* or *I wanna hold your hand*. On seeing the two sample images displayed while the song was playing, one of which was the single's cover, they were instantly able to satisfy their curiosity. Intrigued, they then clicked on the image to explore further.

4.4. Integration

The majority of the components written for the prototype—the plugin, classifier, and format statements—are naturally drawn into Greenstone through a collection's configuration file. To operate with the extended functionality it is merely a matter of ensuring the modules of code are placed in the correct directories.

Integrating the collaging applet into the Greenstone environment required the editing of macro files. This is a mechanism within the toolkit that blends a digital library site's look and feel with the specification of functional aspects of the web page based interface, to precisely cater for this sort of extension. Macro files typically set the global form within the digital library; the aforementioned format statements override this for more local customisations.

With this done a user can create a new collection, set the collection configuration file to use the specialist plugin and classifier, populate it with MP3 files and issue the build command. Greenstone even provides a “base this collection on” feature that means that once a baseline collection of this type has been established, future collections can inherit there configuration settings from this. Then the task becomes: start a new collection, populate it with documents, and then build.

5. CONCLUSION

The development of novel browsing tools in music digital libraries represents a relatively overlooked research domain. Empirical data suggest that browsing should be seen as a fundamental aspect of music information retrieval and music digital library use. To address this shortcoming, we have developed and presented a proof-of-concept prototype that applies collaging, a novel form of digital library browsing, to musical data.

There are many directions in which to extend and develop the work described here. For example, we plan to explore the utility of affording users the ability to “drag and drop” from the presented images to create a user-defined playlist in future versions of this tool. The playlists could actually form the bases for user-personalised sub-collections or mini-libraries drawn from larger, more heterogeneous music digital libraries.

Beyond browsing, but closely related to the personalization functions mentioned above, the music collaging technique can be used in conjunction with traditional search tools. Under this scenario, traditional search tools

could be enhanced to allow users to conduct deliberately “broad” or “fuzzy” searches (i.e., music from the Haight-Ashbury period, horror film music, etc.). The search results from these broad sets would then be presented via the music collaging technique for “laid back” perusal by the user.

We also envision implementations of the music collaging system that are presented on large, touch-sensitive screens located in record stores and music libraries. On the opposite end of the spectrum, the music collaging technique could equally run as a Personal Digital Assistant (PDA) enhancement for users to access their MP3 collections.

6. REFERENCES

- [1] D. Bainbridge, J. Thompson, and I. H. Witten. Assembling and enriching digital library collections. In *Proc. of the Third ACM and IEEE Joint Conference on Digital Libraries*, pages 323–334, Houston, Texas, 2003.
- [2] M. Chang and J. Leggett. Collection understanding through streaming collage. In *Proc. of the Information Visualization Interfaces for Retrieval and Analysis (IVARA) Workshop, associated with the Joint Conference on Digital Libraries*, Houston, Texas, USA, May 2003.
- [3] H. Chen, A. Houston, R. Sewell, and B. Schatz. Internet browsing and searching: User evaluations of category map and concept space techniques. *Journal of the American Society for Information Science*, 49(7):582–603, 1998.
- [4] S. Cunningham, M. Jones, and S. Jones. Organizing music for use: personal music collections. Technical report, Department of Computer Science, University of Waikato, Hamilton, New Zealand, 2004.
- [5] S. Cunningham, N. Reeves, and M. Britland. An ethnographic study of music information seeking: implications for the design of a music digital library. In *Proc. of the Joint Conference on Digital Libraries*, pages 5–16, Houston, Texas, USA, May 2003.
- [6] M. Jones, P. Jain, G. Buchanan, and G. Marsden. Using a mobile device to vary the pace of search. In *International Symposium on Mobile HCI*, pages 8–11, Udine, Italy, September 2003.
- [7] S. Jones, S. Cunningham, R. McNab, and S. Boddie. A transaction log analysis of a digital library. *International Journal on Digital Libraries*, 3(2):152–169, 2000.
- [8] U. Karadkar, A. Kerne, R. Furuta, L. Francisco-Revilla, F. Shipman, and J. Wang. Connecting interface metaphors to support creation of hypermedia collections. In *Proc. European Conf Digital Libraries*, pages 338–349, 2003.
- [9] A. Kerne. CollageMachine: an interactive agent of web recombination. *Leonardo*, 33(5):347–350, 2000.
- [10] A. Kerne, M. Khandelwal, and V. Sundaram. Publishing evolving metadocuments on the web. In *Proc. of the 14th ACM Conference on Hypertext and Hypermedia*, pages 104–105, Nottingham, UK, 2003.
- [11] A. Kerne and V. Sundaram. A recombinant information space. In *Proc. Computational Semiotics in Games and New Media (CoSIGN)*, pages 48–57, 2003.
- [12] J. R. McPherson and D. Bainbridge. Usage of the MELDEX digital music library. In *Proc. of the Second Annual International Symposium on Music Information Retrieval: ISMIR*, pages 19–20, 2001.
- [13] J. Preece, Y. Rogers, H. Sharp, D. Benyon, S. Holland, and T. Carey. *Human-Computer Interaction*. Addison-Wesley, Wokingham, UK, 1994.
- [14] A. Rauber, E. Pampalk, and D. Merkl. Using psycho-acoustic models and self-organizing maps to create a hierarchical structuring of music by sound similarity. In *Proc. of the Third International Conference on Music Information Retrieval: ISMIR*, pages 71–80, 2002.
- [15] G. Tzanetakis, G. Essl, and P. Cook. Automatic musical genre classification of audio signals. In *Proc. of the Second Annual International Symposium on Music Information Retrieval: ISMIR*, pages 205–210, 2001.
- [16] I. H. Witten and D. Bainbridge. *How to build a digital library*. Morgan Kaufmann, San Francisco, 2003.
- [17] I. H. Witten, D. Bainbridge, G. W. Paynter, and S. J. Boddie. Importing documents and metadata into digital libraries: Requirements analysis and an extensible architecture. In *Proceedings of the European Conference on Digital Libraries*, pages 390–405, Sept. 2002.