DIGITAL MUSIC COLLECTIONS: A PLACE OF KNOWLEDGE EMERGENCE?

Francis Rousseaux  Alain Bonardi  Benjamin Roadley
Ircam  Ircam  University of Reims
francis.rousseaux@ircam.fr  alain.bonardi@ircam.fr  ben.roadley@gmail.com

ABSTRACT
Many of our modern computerized activities, may they be personal, professional or even artistic, involve searching, classifying and browsing large numbers of digital objects. The tools we have at hand, however, are poorly adapted as they are often too formal: we illustrate this matter in the first section of this article, with the example of multimedia collections. We then propose a software tool for dealing with digital collections in a less formal manner. Finally, we see that our software design is strongly backed up by both artistic and psychological knowledge concerning the ancient human activity of collecting, which we will see can be described as a metaphor for categorization in which two irreducible cognitive modes are at play: aspexitual similarity and spatiotemporal proximity.

MULTIMEDIA COLLECTIONS
Technological context
Since the early times of WIMP-based interfaces in the 70s, the technology has leaped forward, and today computers are equipped with high storage capacity hard drives, powerful processors, high bandwidth internet connections, to name but a few technological trends. These are still evolving but the fact is that today more and more people are using their computers not only for editing and filing documents, but also for collecting music, films, images, books...
Not surprisingly, a huge market has emerged from these multimedia collections. We can now choose from a myriad of computerized tools which assist us in finding, retrieving, recording, creating, editing, browsing and classifying multimedia contents. The variety of tools at hand seems to fit with the variety of uses involved in multimedia computing, from the most creative ones - such as graphic design, audio synthesis, etc - to the most formal ones – classification in particular. However, there doesn't seem to be many tools bridging the gap between these two seemingly opposing polarities.

Collecting: Between Formalism and Creativity
First, let us suggest that looking for new material and classifying are two important processes involved in collecting. In order to illustrate our point, let us describe a particular example: the music collector. As we have said, our collector will surely possess some initial items; these may be some CDs or vinyl records. His first action involved in extending his collection could be a visit to the record shop for example. Here, the music is classified conformingly to the record companies' desires, which can sometimes be confusing for our collector, who is a fan of Jimi Hendrix, and just does not know where to look for his albums: in the blues section? rock section?
Is there a 'sixties' section? Anyway, despite finding them rather practical at first sight, our collector didn't create these labels, and finds it difficult adapting to them. However, as he browses through the shop, he also notices some nicely illustrated records, and discovers new artists he is interested in because their records are sitting next to Jimi's. Finally, when he has bought enough music records, and come back home, he will be able to start arranging his collection in a very personal and satisfying manner, which will be pleasing to the eyes, and also allow him to retrieve items quickly.
If he had decided to collect digital music, and go online to find new items for his collection, the process would have been rather similar. Commercial music download sites allow the user to browse through predefined music categories, thus implementing a kind of virtual record shop with the same problems mentioned earlier. The search tool however can come in handy, and allow the user to search for the name of an artist, a song, an album or even musical genre. All these are still editorial information, which aren't necessarily the most useful to the collector. Then, when the music is downloaded, the album consists of a group of compressed audio files, containing preset meta-tags, again storing editorial information. When browsing these files in his audio player, the songs are defined and classified automatically, not always according to the collector's desires. His final attempt is then to create a set of folders on his disk, and arrange his items in these folders. But how does he name these folders? What if he wants to arrange and browse the items in multiple ways?
What if a particular item doesn't fit in any folder, or could be placed in two or three different categories? Pachet has also described many problems in the area of Electronic Music Distribution [1].
As we see from this example, the tools that the everyday user has at hand are too formal, and are poorly adapted to the growing activity of collecting multimedia contents.

Attempts have been made at putting the human user back in control of the collecting process, rather than relying purely on predefined categories and automated research algorithms. However, it has become obvious that the other extreme of handing complete control over to the user isn't optimal either, as in online content
sharing sites, like the famous FlickRZ. Therefore, we believe that an optimal solution to the problem of digital collections could lie somewhere between these two polarities: predefined categories and total user creativity.

**Examples of Tools Attempting to Bridge the Gap**

MusicBrowser is a software which aims at indexing large and unknown music collections, and also helping the user find "interesting" music in these collections [2].

When digital sound files are imported into the system, they are analyzed, and a database of their acoustic properties is created / updated. Then the user can browse through the collection in a traditional manner, relying on editorial information. He can also create his own categories intuitively. He starts by creating a category, and giving it a name. This can be totally subjective if he wishes, he may call it "evening music", "happy music" or "favorite", etc. He then adds a few songs to this category, before asking the program to finish classifying, based on acoustic similarities. This creative feedback loop, between user input and automated algorithms will eventually lead to a satisfying classification for the user, who will have saved a lot of time in the process.

IMEDIA is a research project focused on indexing large collections of photos, and interactive searching and browsing [3]. When photos are added to the system, they are analyzed and a database of visual descriptors is created / updated. One of the main features of the program is allowing the user to search for similar photos with a "relevance feedback" system.

In these two systems, we have noticed a creative feedback loop between the human user's input (starting point, examples, relevance feedback...) and the computer (automated algorithms for classifying and searching). This helps the user build and browse his collection in a constructive process, leading to a result which neither he nor the computer could have achieved alone. Also, both editorial information and semantic information (invisible to the user) are taken into account. Furthermore, interactive searching and browsing can be transposed to different media.

We can even think further, and imagine a common environment for collecting multimedia files. This could be a system with a generic layout and set of functionalities that would give birth to different programs specialized in collecting certain types of media. In the next section, we shall present a software prototype that we have implemented in order to experiment with this idea. As we shall see in the next section, we have tried to create a program more suitable to the particular process of collecting, which has an element of subjectivity, evolves over time and doesn't rely purely on similarities.

**THE RECOLLECTION SOFTWARE**

ReCollection is a computer program for searching, arranging and browsing digital content. As our collecting activities vary from one context to another, it is too ambitious to seek a general solution to the problem. Rather, particular application areas must be defined and isolated, in order for a specific answer to be given, however always relying on a set of basic principles. Here, we shall discuss the software prototype we have created for the digital opera / open form opera Alma Sola.

**A Useful Metaphor: the Art Collection**

Artists and philosophers have described some very particular characteristics of collections. One of those, as noted by Wajcman [4], is that of excess in a collection. This means that the number of collected items exceeds the collector's capacity of memorization, but also of physical storage and exposition in the gallery. Thus, there is a need for at least one reserve, where the excess can be stored. Often, the items in reserve are stored in heaps, in random locations, and they aren't always labeled, which makes it difficult to find and retrieve objects. The reserve allows us to handle the excess in collections, which is a problem in many of today's computer applications. On the other hand, objects which are currently exposed are found in the gallery. Here, the objects follow a spatio-temporal arrangement defining a finite number of visitation paths. The closeness in space of certain artworks and the chronological order in which they are approached are set carefully by the curator, as they strongly influence the visitors' experience.

**The Reserve**

The ReCollection software has two main modes: reserve and gallery. The reserve allows us to store our objects which aren't exposed in the gallery. There are many objects in the reserve, and these are not always labeled; also they are rarely arranged in an orderly and tidy manner. So when we visit the reserve, we have no choice but to wander around, picking up objects, inspecting and identifying them one at a time. The reserve can also be compared to the attic, in which our family possessions are stored similarly. As we explore our attic, we can happen to pick up an old photo album, which we had completely forgotten about. This item will surely bring back memories and emotions. We can then choose to keep this album under our arm, as we continue to explore the attic, or we can leave straight away, and put it on our replace, for example, making it visible to visitors. It is all these pleasant and familiar experiences which we believe can be recreated thanks to the modeling of the reserve in our computer program.

![Figure 1. An example of reserve.](http://www.almasola.net)
The user can create any number of reserves. However, he must create at least one, and store at least one object in this reserve. When he is in reserve mode, he can only view one object at a time. When he decides to view another object, it is chosen randomly from the remaining items in reserve. During a visit, each object is viewed only once. If the user wants to view an item he has already visited, he may go through the history of items on the left side of the screen, as shown in Fig. 1. When he finds an object of interest, he can move it to the gallery. It will then be removed from the reserve, and saved in memory, with a group of objects waiting to be imported in the gallery. Then, in gallery mode, the user will see this heap of objects, and will be able to import it in the desired gallery, at the desired location.

The Objects

The items in the Alma Sola collection are made up of three components: a photo of the performance, a sound recording of a few seconds of the singing, a text, the line which is sang in the corresponding sound file. These are all regular files stored on disk (bitmap, wave and .txt formats). In a more general context, the objects can be made up of any one of these types of media, a video (though not implemented in this version), or any combination of these.

Also, each object has a set of descriptors attached. There is a specific set of descriptors for each type of media, which describe the contents of the object, for example the average volume of the sound, the brightness of the photo, the number of words, etc. Depending on the application, we could also include editorial information, such as date, author, etc.

These descriptors may be assimilated to the private properties of traditional computer objects. But in the context of collecting objects, we also need to account for other properties that come from the activities in which these objects collectively engage.

The Gallery

A collective activity involving a number of objects at a time is their relative arrangement in the gallery space. To the location of objects in this space, we have added their color; these two properties make up an extra conceptual layer which is the framework for the creation and management of our collections.

In ReCollection, there is always at least one gallery, and the user can create as many as he wishes. There is always at least one item in a gallery, some basic content that the user can interact with, a starting point for his collection.

The objects can be placed and arranged manually in the gallery space, using click and move, just as in common user interfaces. The user can also rely on two algorithms to automatically dispose the objects. The first one, inspired by cataRT software [5], calculates the objects’ positions and colors according to descriptors chosen by the user. The second calculates the positions depending on a sample of objects selected by the user. A Principal Components Analysis (PCA) finds out which descriptors vary most amongst the objects of the sample, the system can then rearrange the whole gallery according to these descriptors, as in the first method.

The arrangements resulting from the algorithmic calculations can always be modified manually in order to correct them (in the eventuality of rather subjective descriptors), to build up a global figure, or to bring items together.

This way, through creative human-computer feedback loops, meaningful global figures can emerge through the arrangement in space of collected items, as well as local figures, soft pseudo-categories which are heaps of objects brought together by the system and/or the human user. These pseudo-categories are the building blocks for the classes the collection is implicitly aiming for. They are easily and constantly updated; items are added and removed instantly by being moved in space. They are loosely defined and never completely closed off from others, allowing some objects to be lost somewhere in between several heaps, when they cannot be placed in any one category. In a nutshell, this system allows for the creation of collections in which classes are in constant evolution, and are built by exploiting not only the objects’ degree of similarity, but also their relative location in space and time.

Furthermore, the user may wish to search for objects in the gallery or in the reserve, in order to build on these categories, look for new kinds, or even fill in gaps in the gallery space. For this, the ReCollection system has two search tools he can use. The first is a simple 'keyword query', which searches for a keyword within the text or names of the objects. The second is a 'search by similarity'. The user selects an object, or group of objects, and the system searches for items which are similar (according to the descriptors). In both cases, the search is carried out in both the gallery and reserve, and a list of results is displayed in the gallery, ordered by similarity. Once all the items of interest have been imported from the reserve, through browsing or searching, and once they have been arranged in the gallery space, the user has a first disposition he can play with. When he will browse the gallery space, his experience will be influenced by the fact that certain objects are close in space, and in time of visitation. Although this is interesting in itself, the system can help the user go farther, by defining a set of guided visits, which are simply an order of visitation of selected objects in the gallery.

The type of interface we have chosen to implement these functionalities is a 2D zoomable user interface (ZUI), inspired by Ken Perlin’s Pad [6]. All objects are in the same 2D space, which has no borders. The point of view can be moved vertically and horizontally, and the user can zoom in and out. If he zooms in on an item, until it fills the screen, the sound is played back. This kind of interface has been experimented; it has obtained good results, and has been proven reliable [7]. Its intuitive approach is seducing to us, particularly in our goal of intuitively collecting digital media. Finally, the spatial metaphor takes advantage of the users’ spatial memory and cognitive abilities [8, 9].
Husserl used to say that consciousness is always conscious of something, that consciousness always pre-dates the subject and the object, and puts them together in the process. There are no subjects or objects already existing independently that meet in the world to fill out a journal of experiences (the subject) and perhaps adapt to each other by induction. In the same fashion, we could say that a collection is always a collection of something, in that the original process of categorization is the activity of collecting, implacably mixing abstraction and spatio-temporal arrangements, and producing as many metastable categories.

The current models for information search are too formal, and they assume that the function and variables de ning the categorization are known in advance. In practice, however, when searching for information, experimentation plays a good part in the activity, not due to technological limits, but because the searcher does not know all the parameters of the class he wants to create. He has hints, but these evolve as he sees the results of his search. The procedure is dynamic, but not totally random, and this is where the collection metaphor is interesting.

The collector's experimentation is always carried out by placing objects in temporary and metastable space/time. Here, the intensity of the future category has an extensive nature in space/time. And this system of extension (the figure) gives as many ideas as it does constraints. What is remarkable is that when we collect something, we always have the choice between two systems of constraints, irreducible one to the other. This artificial indi erentiation for similarity/contiguity is the only possible kind of freedom allowing us to categorize by experimentation. Our prototype implements these ideas by allowing the user to dispose his objects in 2D space. This arrangement may be manual, automated or both; it may be based on similarity, spatial proximity or both. A global figure may emerge from this arrangement, in uencing the browsing and also the extension of the collection. Local figures emerge, which are the temporary pseudo-classes illustrating the pre-categorization building process of collecting. The art gallery metaphor fits very well, as it adds further meaning to the arrangement of the collected items in space, and models the excess in collections thanks to the reserve.

Through exploiting space in this way, the software interface takes advantage of our cognitive abilities in dealing with spatial information, and also our ability to collect information and acquire knowledge. Our next step is experimentation in order to validate our work. This could simply take the form of a series of sessions in which both novice and experimented users are asked to build up collections using the software. Through user-feedback, we will have a first idea of how well the interface is understood, how useful the users find it and how easy it is to use. If this experiment is a success, as we believe it will be, we will continue our research and bring it to the next level. Through integrating new functionality focused on indi erentiation for similarity/proximity, we will be able to build specific tools for a variety of applications in which the user's activity may be - at least metaphorically - described as building a figural collection.

REFERENCES


