Exploratory Browsing: Enhancing the Browsing Experience with Media Collections

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik

der Ludwig-Maximilians-Universität München

von

Yaxi Chen

München, den 10. Juni 2010

Exploratory Browsing: Enhancing the Browsing Experience with Media Collections

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik

der Ludwig-Maximilians-Universität München

von

Yaxi Chen

München, den 10. Juni 2010

Berichterstatter:

Prof. Dr. Andreas Butz
(Ludwig-Maximilians-Universität München, Deutschland)

Prof. Dr. Susanne Boll
(Universität Oldenburg, Deutschland)

Tag der mündlichen Prüfung: 20. Juli 2010

Abstract

In recent years the digital media has influenced many areas of our life. The transition from analogue to digital has substantially changed our ways of dealing with media collections. Today's interfaces for managing digital media mainly offer fixed linear models corresponding to the underlying technical concepts (folders, events, albums, etc.), or the metaphors borrowed from the analogue counterparts (e.g., stacks, film rolls). However, people's mental interpretations of their media collections often go beyond the scope of linear scan. Besides explicit search with specific goals, current interfaces can not sufficiently support the explorative and often non-linear behavior. This dissertation presents an exploration of interface design to enhance the browsing experience with media collections. The main outcome of this thesis is a new model of *Exploratory Browsing* to guide the design of interfaces to support the full range of browsing activities, especially the *Exploratory Browsing*.

We define *Exploratory Browsing* as the behavior when the user is uncertain about her or his targets and needs to discover areas of interest *(exploratory*), in which she or he can explore in detail and possibly find some acceptable items (*browsing*). According to the browsing objectives, we group browsing activities into three categories: *Search Browsing*, *General Purpose Browsing* and *Serendipitous Browsing*. In the context of this thesis, *Exploratory Browsing* refers to the latter two browsing activities, which goes beyond explicit search with specific objectives.

We systematically explore the design space of interfaces to support the *Exploratory Browsing* experience. Applying the methodology of User-Centered Design, we develop eight prototypes, covering two main usage contexts of browsing with personal collections and in online communities. The main studied media types are photographs and music.

The main contribution of this thesis lies in deepening the understanding of how people's exploratory behavior has an impact on the interface design. This thesis contributes to the field of interface design for media collections in several aspects. With the goal to inform the interface design to support the *Exploratory Browsing* experience with media collections, we present a model of *Exploratory Browsing*, covering the full range of exploratory activities around media collections. We investigate this model in different usage contexts and develop eight prototypes. The substantial implications gathered during the development and evaluation of these prototypes inform the further refinement of our model: We uncover the underlying transitional relations between browsing activities and discover several stimulators to encourage a fluid and effective activity transition. Based on this model, we propose a catalogue of general interface characteristics, and employ this catalogue as criteria to analyze the effectiveness of our prototypes. We also present several general suggestions for designing interfaces for media collections.

Zusammenfassung

In letzter Zeit haben digitale Medien bereits viele Bereiche unseres täglichen Lebens beeinflusst. Der Übergang von Analog zu Digital hat die Art der Medienbehandlung wesentlich geändert. Die bestehenden Interfaces für digitale Medien bietet hauptsächlich feststehende lineare Modelle an. Solche Modelle folgen entweder zu Grunde liegenden technischen Konzepten (Ordner, Veranstaltungen, Alben, etc.) oder Metaphern ihrer analogen Komponenten (z.B. Stapel, Filmrollen). Die mentale Interpretation von Mediensammlungen geht jedoch häufig über den Bereich eines linearen Scans hinaus. Neben expliziter Such mit spezifischen Zielen, können bestehende Interfaces das explorative Verhalten, das oft nicht linear ist, nicht ausreichend unterstützen. Diese Dissertation stellt eine Exploration für Interface Design vor, um das Browsing Erlebnis in Mediensammlungen zu verstärken. Das Resultat der Dissertation ist ein neues Modell für *Exploratory Browsing*, mit dem das Interface Design besser gesteuert und die Aktivitäten des Browsings, besonders *Exploratory Browsing*, unterstützt werden können.

Wir definieren *Exploratory Browsing* als das Verhalten, wenn der Benutzer sich über das Ziel im Unklaren ist während er einen interessanten Bereich herausfinden möchte (*exploratory*). Danach kann der Benutzer den ausgewählten Bereich erkunden und möglicherweise akzeptable Elemente finden (*browsing*). Wir gruppieren die Browsing Aktivitäten in drei Kategorien (basierend auf den Zielen des Browsings): *Search Browsing*, *General Purpose Browsing* und *Serendipitous Browsing*. Im Kontext dieser Dissertation bezieht sich *Exploratory Browsing* auf die letzten zwei Kategorien, die unterschiedlich von expliziter Suche mit speziefischen Zielen sind.

Um das Erlebnis des *Exploratory Browsings* zu unterstützen, erforschen wir systematisch den Design Space der Interfaces. Mit der Methode des User-Centered Designs haben wir acht Prototypen entwickelt. Diese konzentrieren sich auf zwei Anwendungsaspekte – private Mediensammlungen und online Gemeinden. Die verwendeten Medienntypen sind Fotos und Musik.

Der Hauptbeitrag dieser Dissertation liegt darin, das Verständnis zu vertiefen, wie sich exploratives Verhalten der Menschen auf das Interface Design auswirkt. Die Dissertation hat zu dem Forschungsbereich des Interface Designs in den folgenden Aspekten beigetragen. Um das Ziel zu erreichen, dass *Exploratory Browsing* in Mediensammlungen im Interface Design unterstützt werden kann, präsentieren wir ein Modell für *Exploratory Browsing*. Das Modell beinhaltet alle explorativen Aktivitäten der Menschen in Mediensammlungen. Des Weiteren untersuchen wir das Modell in verschieden Anwendungsbereichen und entwickeln acht Prototypen. Die umfangreichen Erkenntnisse und Implikationen, die in der Entwicklung und Auswertung der Prototypen gesammelt wurden, werden dann zur Verbesserung unseres Modells genutzt: Wir decken die Relationen zwischen Browsing Aktivitäten auf. Gleichzeitig finden wir einige effektive Stimulatoren, mit denen wir eine fließende und effektive Übergabeaktivität garantieren können. Basierend auf dem Modell stellen wir einen Katalog für die Eigenschaften der allgemeinen Interfaces vor. Wir benutzen den Katalog als Kriterium, um die Effektivität unserer Prototypen auszuwerten. Ferner geben wir mehrere allgemeine Vorschläge zu Interface Design für Medienasammlungen.

Acknowledgements

As of today, when I come to the end of my PhD, there are a lot of people I deeply appreciate for their support, encouragement and inspirations. I'll never forget these people who have helped me in the past three years. I sincerely apologize if I fail to mention someone.

First of all I deeply thank my advisor Prof. Andreas Butz, who provided me this research opportunity and guided me into the new world of Human-Computer Interaction. He walked me through the whole process, listened to me and inspired me in each conversation between us. He also helped me avoid any unnecessary troubles or risks in my work. A deep thanks goes to Prof. Susanne Boll for investing her valuable time and effort into helping me refine the structure of this thesis. I also thank Prof. Hussmann and Prof. Ohlbach for their willingness to be on my committee.

During my expedition, I fortunately had the opportunities to get to know some excellent researchers from different fields. They are one of my key sources of inspirations. I want to thank Heidi Lam for helping me coming up the structure of the thesis. I also want to thank Rodrigo Santamaría and Therón Roberto, working with whom enriched my experience of collaboration. Thank Rodrigo for always encouraging me during the thesis writing.

I felt very lucky to have worked in an energetic and fruitful research environment. Thank my competent and creative colleagues in the group of Media Informatics for their company. Otmar Hilliges and Sebastian Boring are always the right people to talk about ideas, papers and even troubles. I'd like to thank Dominikus Baur and Michael Sedlmair. Discussions with them provided me valuable feedback and inspirations. A special thank goes to our „Wanna Be PhD" (WBP) workshop: Thank Richard Atterer for his encouragement especially in the orientation phase of my PhD. Thank Florence Balagtas-Fernandez and Sara Streng for always listening to me and sharing experience with me. I sincerely thank the students I worked with for their creativity and contributions: Franz Berwein, Matthias Hoyer, Susanne Keck, René Klüber, Christian Köpke, Florian Lambers, Silvia Peter, Michael Reiter and Michael Schmidt.

I truly thank the China Scholarship Council (CSC) and the German state of Bavaria for the financial support. I also want to thank the warmhearted staffs in the international office at the University of Munich. Thank my friends for their emotional support and the unforgettable time we spent together in Oberschleißheim.

I specifically thank my family for their selfless love and unconditional support. My parents have spent so much time and energy on me since I was a little child. They showed me good examples in how to become a diligent, righteous and modest person. Thank my parents in law for taking care of my daily life in Chengdu.

Finally I want to thank the most special person in my life: my husband Yangchun. Without your support and encouragement I cannot reach where I am today. I hope we will continue to support each other and experience the adventure together in the rest of our life.

Contents

I Introduction and Motivation

List of Figures

Part I

Introduction and Motivation

This thesis is structured into four parts. Each part is built based on the previous ones, which illustrate the evolution of the process and methodology experienced in this research.

This part forms the foundation of this research. Chapter 1 introduces the motivation, research context and methodology of this thesis. It also presents our definition and initial model of *Exploratory Browsing*. We argue the necessity to inform a new model to guide the interface design to support the full range of browsing activities, especially the *Exploratory Browsing* with media collections. Chapter 2 presents an overview of related work on interface design for media collections. This discussion functions as the fundamental for our exploration in the subsequent parts.

Part II and III present our investigation to support the *Exploratory Browsing* experience in two main usage contexts respectively: with personal media collections and in online communities. Finally Part IV summarizes the process, methodology and outcomes of this research. It concludes this dissertation with discussions on future work.

1

Introduction

This thesis explores the interface design for media collections, with the ultimate goal to inform a new model for designing interfaces to support the exploratory experience with media collections. This chapter introduces the motivation, research context and methodology which form the foundation of this dissertation. It introduces our definition and initial model of *Exploratory Browsing*. It then discusses the main contributions and the structure of this thesis.

1.1 Motivation

The digitalization of media has influenced many areas of our life, and in particular it has changed our ways of dealing with media collections. Photos and CDs are now represented as bits and bytes. Their physical containers of albums, boxes and shelves, are replaced by virtual icons and folders (see Figure 1.1c and d). Not only the representation and archive forms, this digitalization is changing our behavior with media collections. Some ways of dealing with our analogue collections, which we have learned naturally lifelong, are not suitable for the digital counterparts anymore, such as paging through a photo album (see Figure 1.1a), or searching for a CD placed in one corner of the shelf (see Figure 1.1b). Besides this affordance inconsistency between analogue and digital, we still receive tremendous benefits with digital media: We can take a series of photos of a certain object and view them right after shooting; we may put a lot of music into our mp3 player and play it in a shuffle mode during commute.

Concerning software for managing these media collections, today's interfaces mainly offer linear models corresponding to the underlying technical concepts (e.g., folders, events, albums), or the metaphors borrowed from the analogue counterparts (stacks, film rolls, etc.). Existing interfaces for photo collections (e.g., Picasa[1]) often use the camera's time stamps to determine

[1] http://picasa.google.com

events, i.e., groups of photos which were shot closely after each other. iPhoto[2], for example, employs the analogue metaphor of a film roll to group photos which are read out in a batch. It has recently become a fashion to display photos on a map based on their geo-coordinates, which is already applied in iPhoto and Picasa. Current interfaces for music collections almost exclusively use the conceptual hierarchy of artist-album-song, which is currently still the mainstream way to publish music. Additional information, such as genre, year, tempo, etc., is used for sorting and filtering, and the collection or search result is always displayed in a form of list (in some applications such as iTunes[3] enhanced with cover arts).

(a) (b) (c) (d)

Figure 1.1: Analogue media collections and their digital counterparts: a) Page through a photo album (Peter, 2010). b) Look for an album in one corner of the shelf (Klüber, 2009). c) Navigate the photo folders in Windows Explorer[4]. d) Scroll a long playlist in iTunes[3].

These interfaces support certain activities very well, which cannot easily be achieved with analogues, for example, reorganizing collections along different criteria, or quickly locating a specific known item within a large set, which used to be a tedious work with large analog

[2] http://www.apple.com/ilife/iphoto

[3] http://www.apple.com/itunes

[4] http://www.microsoft.com/windows/windows-vista/features/explorers.aspx

collections. However, other important activities around media collections are more dynamic and often unstructured, such as casual browsing and searching, or loosely structured story-telling (Balabanović et al., 2000; van House et al., 2004). People often have no well-defined needs and rarely look for one specific item. Most of their time is spent in browsing-alike activities and they may identify the match “when they see it”. People attempt to find an acceptable alternative rather than the optimal solution. They often change their mind and easily get sidetracked by unexpected but interesting findings. This exploratory behavior and flexible association with media collections are poorly supported by today’s interfaces. We believe that applications supporting these exploratory activities will be appealing and lead to more powerful interfaces.

The main focus of this thesis is to enhance the browsing experience with media collections. Through a systematic investigation of the full range of browsing activities in different usage contexts, we seek for a new model of *Exploratory Browsing*, which can guide the design of interfaces to support the *Exploratory Browsing* experience with media collections.

1.2 Exploratory Browsing

In this thesis, we define *Exploratory Browsing* as the behavior when the user is uncertain about her or his targets and needs to discover areas of interest (*exploratory*), in which she or he can explore in detail and possibly find some acceptable items (*browsing*). According to the browsing objectives, we group browsing activities into three categories: *Search Browsing*, directed and structured activity with specific goal; *General Purpose Browsing*, exploration in a certain area of interest with unspecified objective; *Serendipitous Browsing*, unstructured and random activity. In this thesis, *Exploratory Browsing* refers to the latter two activities which involve no specific search goal.

This section presents our initial model of *Exploratory Browsing*, which is inspired by the existing models of searching and browsing. The mental model is first introduced, followed by the model of Exploratory Search in the field of Information Retrieval, in which the full range of search activities is well categorized. This model particularly highlights the importance of satisfying people’s exploratory behavior beyond the query-based search, which has also been investigated in other research. Based on these existing works, we present our initial model of *Exploratory Browsing* in the context of media collections.

1.2.1 Mental Model

A mental model is a cognitive representation of information, which is important for exploring an unknown information space (Marchionini and Shneiderman, 1988). The basic mental model is categorized as vertical and lateral thinking. Vertical thinking is the logical, utilitarian thought which "looks at a reasonable view of a problem or situation and works through it, generally in a path of least resistance" (De Bono, 1990). Vertical thinking is goal-oriented, selective and straightforward, which concentrates on the relevance while excluding the irrelevance. Conversely, lateral thinking is defined in Oxford English dictionary as "[…] a way of thinking which seeks the solution to intractable problems through unorthodox methods, or elements which would normally be ignored by logical thinking." It is a generative and indirect process, in which breaks and jumps may appear, induced by accidental discoveries (De Bono, 1990). Lateral thinking is prominent when pure browsing activities are carried out (Campos and de Figueiredo, 2001).

1.2.2 Exploratory Search

Regarding the model of lateral thinking, the exploratory aspects of search activities have drawn research attention. Oxford English dictionary terms search as "examination […] for the purpose of finding a person or thing […] investigation of a question; effort to ascertain something." Nowadays, The Web has become one of the main sources for learning and knowledge acquisition. When we meet problems which we have not enough knowledge to solve, we tend to use the Web to look for answers. For example, consider the following three tasks a person might conduct in the Web:

1. Check the time difference between Munich and Beijing.
2. Look for some inspirations for the kitchen decoration.
3. Decide on a place to go for the next trip.

The first task can be effectively solved by a standard query. The user can enter keywords or a short phrase (for example, "time difference between Munich and Beijing") and the system will simply list the relevant answers. For the second task, a list of results is not sufficient to answer such an open question. The searcher tends to define (and possibly refine) the query, look through different sets of retrieval results, and might be attracted by some items during this process. The third task to decide on a tourism destination may need even more effort for information processing, such as comparison, analysis and comprehension, in order to obtain enough knowledge to make the final decision.

These examples illustrate that questions with concrete answers can be efficiently solved through searching with well-defined keywords. However, people's information needs often go beyond explicit query-based search, varying from concept comprehension to lifelong experience accumulation (Marchionini, 2006). In the latter two situations, people may have difficulties to express their needs precisely and verbally, thus keywords are not sufficient to return results matching the high-level expectations. In most of the cases, people have to navigate through the retrieval results, filter the relevant information, compare and analyze it, and then synthesize answers themselves.

In the field of Information Retrieval, researchers have already realized the limitation of the prevalent query-based search, especially in situations when users have no efficient knowledge or contextual awareness to formulate queries (Janiszewski, 1998; White et al., 2006). Systematically, Marchionini (2006) presented the model of Exploratory Search, in which search is represented as a three-stage process of lookup, learning and investigation, and the latter two constitute the Exploratory Search (see Figure 1.2).

Lookup is the most common search activity supported by the conventional search engines. Well-defined queries yield precise retrieval results, which are mostly discrete facts such as numbers, names or specific files. Lookup is also termed as fact retrieval, known item search or question answering. In learning search, time is mostly spent on exploring in a neighborhood of interest and deriving underlying relationships between items through comparison, comprehension and interpretation. More complex activities such as analysis, synthesis and evaluation, are involved in investigative search, a rather long period, to help fostering new knowledge (Marchionini, 2006). The whole search process may be iterative, and the three search activities are thus interpreted in Figure 1.2 as overlapping clouds, as they may intersect with each other.

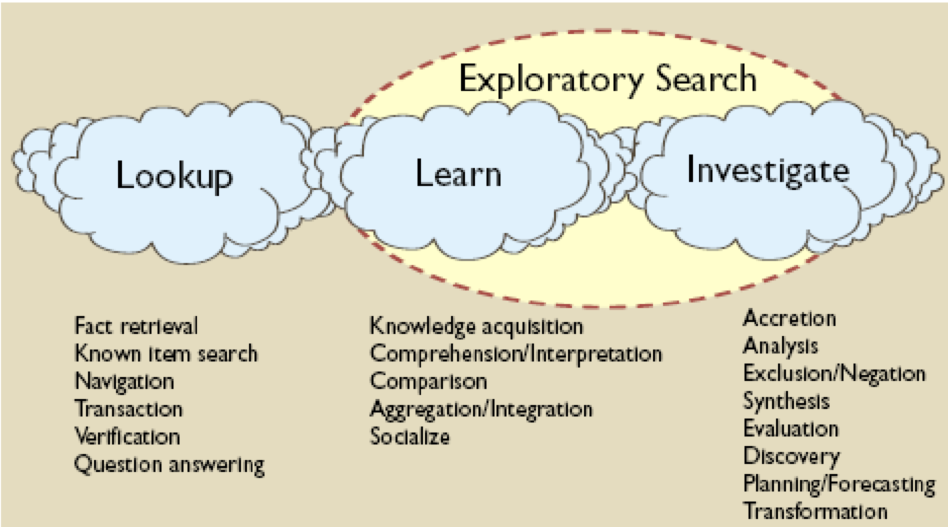


Figure 1.2: The model of Exploratory Search (Marchionini, 2006).

Concerning the interaction between the user and the system, lookup is a turn-taking process, in which both sides take the control in turn. People tend to be more actively involved in the exploratory process. Therefore, to ensure a more fluid exploration, Marchionini (2006) suggested tightly integrating users into the Exploratory Search process through highly interactive interfaces. Ahlberg and Shneiderman (1994a) presented the concept of Dynamic Query, in which the visualization of retrieval results is dynamically updated according to the user's actions, such as adjusting parameter through a slider. HomeFinder (Williamson and Shneiderman, 1992) and FilmFinder (Ahlberg and Shneiderman, 1994b) are two representative examples of such concept. Some research supports moving through large information spaces in flexible ways, such as facet-based investigation in mSpace (Schraefel et al., 2005) and Flamenco (Hearst, 2002; Yee et al., 2003).

1.2.3 Exploratory Browsing in Media Collections

Marchionini and Shneiderman (1988) differentiated search and browsing by the focus of the task. Search is more directed and target-oriented, while browsing is defined as “an exploratory and information-seeking strategy that depends upon serendipity” (Marchionini and Shneiderman, 1988). Cove and Walsh (1988) defined browsing as the actions of selecting useful data from a given information space. Since recognition on sight is easier than pure mental recall, browsing is further described as “the art of not knowing that one wants until one finds it”. This definition also highlights the exploratory aspect of people’s information seeking behavior, which is especially appropriate for solving ill-defined problems and exploring unfamiliar domains. People’s information seeking behavior has been investigated in the field of consumer psychology, which addresses the necessity to focus not only on directed search but also on non-directed exploratory activities (Baumgartner and steenkamp, 1996; Moe, 2007). Users prefer a browsing-oriented interface for browsing task and another interface for direct search with precise targets (English et al., 2002; Hearst et al., 2002).

Lam (2005) proposed the concept of *Exploratory Browsing*, aiming at facilitating search without well-defined targets. *Exploratory Browsing* is described as the behavior when the user is uncertain about her or his targets and needs to define or refine them with multiple related queries. Specifically, she defined exploratory as discovering a local neighborhood of interest, and browsing as exploring this area of interest in detail. Query refinement is enhanced by visualizing individual queries, with the assistance of the history of all the previous queries and a list of suggested query terms. *Exploratory Browsing* can make the exploration flexible and enjoyable, and stimulate inspiration and creativity (Bauernfeind and Zins, 2006).

In the context of media collections, people seldom search for one specific item. They may have no specific objective and attempt to find acceptable alternatives rather than the optimal solution. This kind of exploratory experience (which is discussed in detail in Chapter 2) cannot be sufficiently satisfied by query-based search. We agree on Lam’s understanding of exploratory and browsing. We adapt her definition of *Exploratory Browsing* in the context of media collections and extend it as the behavior when the user is uncertain about her or his targets and needs to discover areas of interest (exploratory), in which she or he can explore in detail and possibly find some acceptable items (browsing).

Based on the definition of *Exploratory Browsing* with media collections, we seek more detailed representations of the relevant activities.

Toms (2000a) describes three typical ways of information acquisition: Seeking a well-defined and known object, or an object that cannot be fully described, but will be recognized on sight, and acquiring information of an object in an accidental or serendipitous discovery.

Vickery (1997) proposed three browsing categories:

- Purposive browsing: Looking for certain information in a defined neighborhood.
- Capricious browsing: Random examination without definite purpose.
- Exploratory Browsing: Searching for inspirations.

Similarly, Cove and Walsh (1988) also identified three browsing activities:

- Search browsing: Directed and structured search for known target.
- General purpose browsing: Exploration within an area of interest.
- Serendipity browsing: Purely random and unstructured activity.

Taking the existing models of exploratory search (Marchionini, 2006) and browsing (Vickery, 1997; Cove and Walsh, 1985) as references, we present the initial model of *Exploratory Browsing* (see Figure 1.3), which consisted of three activities: *Search Browsing*, *General Purpose Browsing* and *serendipitous browsing*. In the context of this thesis, *Exploratory Browsing* refers to the latter two browsing activities, which go beyond explicit search with specific targets. Similar to Exploratory Search (see Figure 1.2), these browsing activities are conceived as interconnected, as people may engage in multiple activities in parallel or switch between different tasks dynamically.

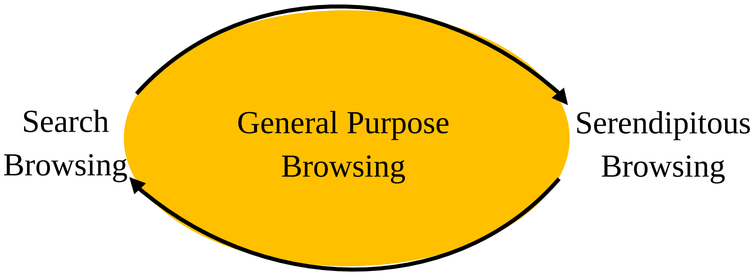


Figure 1.3: The initial model of Exploratory Browsing.

With media collections, *search browsing* is structured and straightforward, triggered by a well-defined objective. If the location of the item is known, the user can navigate through the collection to locate it. With known item name, the user can conduct a query and precise term will yield retrieval results with high relevance, in which the user can easily find the target.

People seldom look for one specific item and instead often browse through their collections (see Chapter 2 for more details), for example, to look for a nice family picture or to compile a playlist for a party. When the user has no clear goal which can be precisely and verbally defined by keywords, *General Purpose Browsing* appears: The loosely defined queries may lead to a broader relevant neighborhood, in which the user puts more effort in comparison, analysis and understanding within different relevant facets. This possibly iterative process can help the user to gradually formulate her or his requirements or identify the item on the way of browsing.

During search and *General Purpose Browsing*, the user's successive actions convey consistent conceptual or semantic relation and they share one general purpose to find the desired information (Toms, 2000b). In these two cases, the user may get sidetracked by unexpected but interesting findings. For example, a user wants to select a photo to print on a Christmas postcard. She starts with some landscape photos. After browsing for a while, she realizes that a photo would look nicer, if all the family members appeared in. Then she changes her mind and finally selects one of those photos. Unexpected discovery leads to *Serendipitous Browsing*. Serendipity is defined in the Oxford English Dictionary as “the faculty of making happy and unexpected discoveries by accident”. Serendipitous discoveries can stimulate curiosity and thus encourage exploration (Roberts, 1989; Toms, 2006b). Aimless browsing is an even more flexible activity, which we group in the scope of *Serendipitous Browsing*. Aimless exploration occurs when a user starts browsing with no priori intentions and let the interests triggered by chance. Without explicit goal, the user may concern more in coverage and exploration (Toms, 2006b). Some researchers argued that *Serendipitous Browsing* is more productive than people initially expect (Callery, 1996). Since it is impossible to program serendipity, the key concern of *Serendipitous Browsing* is to induce and facilitate serendipity (Campos and de Figueiredo, 2001; Toms, 2006b).

Search Browsing suspends when the pre-defined target is found. On the contrary, open-ended *Exploratory Browsing* often produces no concrete output or optimal solution, and instead emphasizes more on the process. As Lam (2005) claimed, “getting there is half of the fun.” Current interfaces for organizing and browsing media collections largely rely on the hierarchical structure or time stamps, both of which offer a fundamental browsing style of linear scan. However, people’s mental interpretation of their media collections often goes beyond this linear navigation. The fixed pre-defined structure makes the *Exploratory Browsing* less fruitful. In this thesis we focus on enhancing the *Exploratory Browsing* experience with media collections. We systematically investigate the model of *Exploratory Browsing*, specifically the transitional relations between browsing activities in different usage contexts. The findings achieved in this exploration help to refine the model of *Exploratory Browsing*.

1.3 Research Context

As shown in Figure 1.4, this thesis locates in a multidisciplinary field of Human-Computer Interaction (HCI), Information Retrieval (IR) and Information Visualization (InfoVis, specifically Casual Information Visualization). The former two areas have achieved tremendous research attention and Casual InfoVis is a rather new research field.

Card et al. (1999) defined InfoVis broadly as computer user uses visual representations to interactively amplify cognition. Pousman et al. (2007) extended the realm of InfoVis from working environment to daily life. This Casual InfoVis helps to “depict personally meaningful information in visual ways that support everyday users in both work and non-work situations”. Casual InfoVis differs from traditional InfoVis in four aspects: The user population is extended from expert to novice; the usage scenario expands from work to daily life; the work-relevant data type is enriched by personal information; gained insights are not analytical, rather reflection of personal and social status. Pousman et al (2006) introduced three main categories of Casual InfoVis: In Ambient InfoVis, interaction is restricted and data is mostly shown in its raw form (Pousman et al., 2006; Skog et al., 2003). For example, InfoCanvas (Stasko et al., 2004) displays electronic paintings, in which certain objects represent the status of corresponding personal information. Social InfoVis is visualization about social network, such as an overview of an online community in PeopleGarden (Xiong and Donath, 1999) and Vizster (Heer and Boyd, 2005), and a visualization of email content in Themail (Viégas et al., 2006). Artistic InfoVis deals mainly with art works (Viégas and Watenberg, 2007), such as visualization of personal activities (Romero and Mateas, 2005) or traffic flow in a public place (Viégas et al., 2004). Besides these three main categories, there are some edge cases, such as Personal InfoVis, which copes with personal data, for example, visualizing personal photo collection in Photomesa (Bederson, 2001) or music listening history in Last.fm Explorer (Pretzlav, 2008).

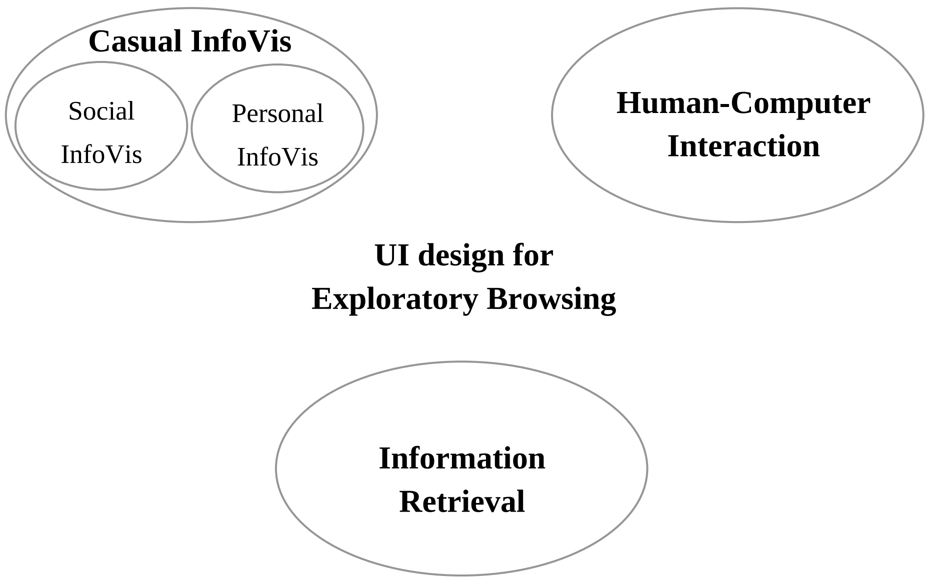


Figure 1.4: The research context of this thesis.

Designing interfaces to support exploratory behavior is a challenging research question. Our exploration in this thesis started with personal media collections, which was later extended to the contexts of online communities. Personal and social aspects are our main concerns in the scope of InfoVis. This thesis is an exploration of designing interfaces to enhance the *Exploratory Browsing* experience with media collections, and we have confronted with the following challenges in our research:

Browsing patterns: Different from typical query-based search, one challenge for this research is to identify the scenarios when *Search Browsing* is not sufficient to satisfy the users' needs. To cover the full range of browsing activities and to ensure an effective transition between them, the browsing patterns need thorough investigation. It is essential to gain a better understanding of the situations, when and how one browsing activity transforms into another.

Interface design: To support explorative activities, a high degree of flexibility is necessary, and users should be allowed to explore the collections in a flexible and intuitive way. The system's representations should be multi-optional and adjustable according to the users' dynamically changing requirements. User interactions should be tightly coupled with the system intelligence, to ensure a fluid and efficient browsing experience.

Evaluation: The system effectiveness is normally evaluated in an usability test, in which quantitative data such as completion time and error rates are measured. Assigning specific tasks may be helpful to achieve comparable quantitative results, but will possibly keep users away from practical usages. User satisfaction is only partially determined by the system efficiency or productivity, and cannot be efficiently reflected by quantitative metrics. Therefore, the users' experience during the exploration should be also considered in the evaluation (Wilson, 2007; Hearst et al., 2002).

1.4 Research Methodology

The exploration in this thesis follows the methodology of User-Centered Design (Norman and Draper, 1986; Preece et al., 2001; Shneiderman and Plaisant, 2005), the research approach widely applied in the filed of HCI. To achieve appropriate outcomes, end users should be integrated in each stage of the design process. Researchers need closely cooperation with users, analyze and understand their requirements. User-Centered Design is normally described as an iterative three-phase process (Sharp et al., 2007):

Analysis: Users, their environment and tasks need thorough investigation. Since user variations may lead to different requirements (Ware, 2004; Shneiderman, 2000; Wijk, 2006), the target user group needs careful identification. Both quantitative and qualitative methods help data collection: Questionnaire is often used to gather quantitative feedback from a large audience. Interview and user observation are the main methods to collect detailed qualitative information.

Design: Based on the insights gained in the analysis phase, different design concepts can be developed. They should be evaluated in their early stage, in order to examine whether they meet the users' requirements. During the assessment, these initial concepts are normally represented as prototypes (Sharp et al., 2007): Using low fidelity prototypes of sketch, storyboard and paper prototype, is the easy and fast way to design and refine interfaces. High fidelity prototypes, such as video prototypes, require more effort to create an elaborate simulation of the interface.

Evaluation: During the evaluation of the fully implemented design, both the system usability (Lauesen, 2005; Shneiderman and Plaisant, 2005) and the users' experience should be investigated to provide a thorough assessment of the system effectiveness (Graham et al., 2000; Spence, 2001). Field study in a real environment further helps to evaluate the suitability of the system in a practical context.

This user-centered approach has already drawn enormous attention and been widely applied in the field of interface design. In this thesis we illustrate how to apply this strategy in designing interfaces for media collections. Figure 1.5 illustrates the design evolution of one interface concept.

(a) (b) (c)

Figure 1.5: Examples of developing a magnet metaphor for photo searching: a) Collecting initial idea through sketching (Reiter, 2009). b) Illustrating the idea with paper prototype (Reiter, 2009). c) The fully implemented concept.

1.5 Contribution

The main contribution of this thesis lies in the understanding of how people's exploratory behavior has an impact on the interface design. The ultimate goal is informing how to afford the model of *Exploratory Browsing* as basis for designing interfaces to support exploratory activities with media collections.

This thesis contributes to the interface design for media collections in several aspects. We present a model of *Exploratory Browsing*, which reveals the full range of exploratory activities around media collections. We investigate this model in different usage contexts, which helps to systematically explore the design space of *Exploratory Browsing* interfaces. During the development and evaluation of eight prototypes, we illustrate how to apply the user-centered strategies in interface design for media collections. The major contributions of this thesis are:

A new model of Exploratory Browsing: We propose a new model of *Exploratory Browsing*, aiming at guiding the interface design to support the highly flexible and explorative behavior around media collections. We systematically explore three main browsing activities: *Search Browsing*, *General Purpose Browsing* and *Serendipitous Browsing*. During our investigation, we uncover the underlying transitions between these activities and discover several stimulators to encourage a fluid and effective activity transition. Based on this model, we propose a catalogue of general characteristics for such a browsing interface. This catalogue is afforded as criteria to analyze the effectiveness of our prototypes. We also propose several suggestions for designing browsing interfaces for media collections.

Exploration of usage contexts: We explore the usage scenarios of *Exploratory Browsing* in two main contexts: browsing with personal collections and in online communities. We systematically examine the model of *Exploratory Browsing* in each context, in order to gain a better understanding of the transition between browsing activities under different contexts. By addressing different personal and collaborative aspects, this research provides an opportunity to investigate and reflect on user needs in diverse domains. This investigation in turn informs the model of *Exploratory Browsing* the fundamental functionalities and tasks in each context.

Research methodology for interface design: This thesis illustrates how to apply the user-centered strategy in designing interfaces for media collections. Users are closely involved in the entire process of analysis, design and evaluation. We develop and evaluate eight prototypes. During the evaluation, we apply different evaluation approaches, considering different characteristics of these systems. Users' interpretations of media collections are mostly personal, and thus lead to personalized outcomes. We investigate the quality measurement of such an output and present several suggestions for its quantitative and qualitative measurements. The implications gathered in the development and evaluation of these prototypes inform the refinement of the initial model of *Exploratory Browsing*.

1.6 Thesis Structure

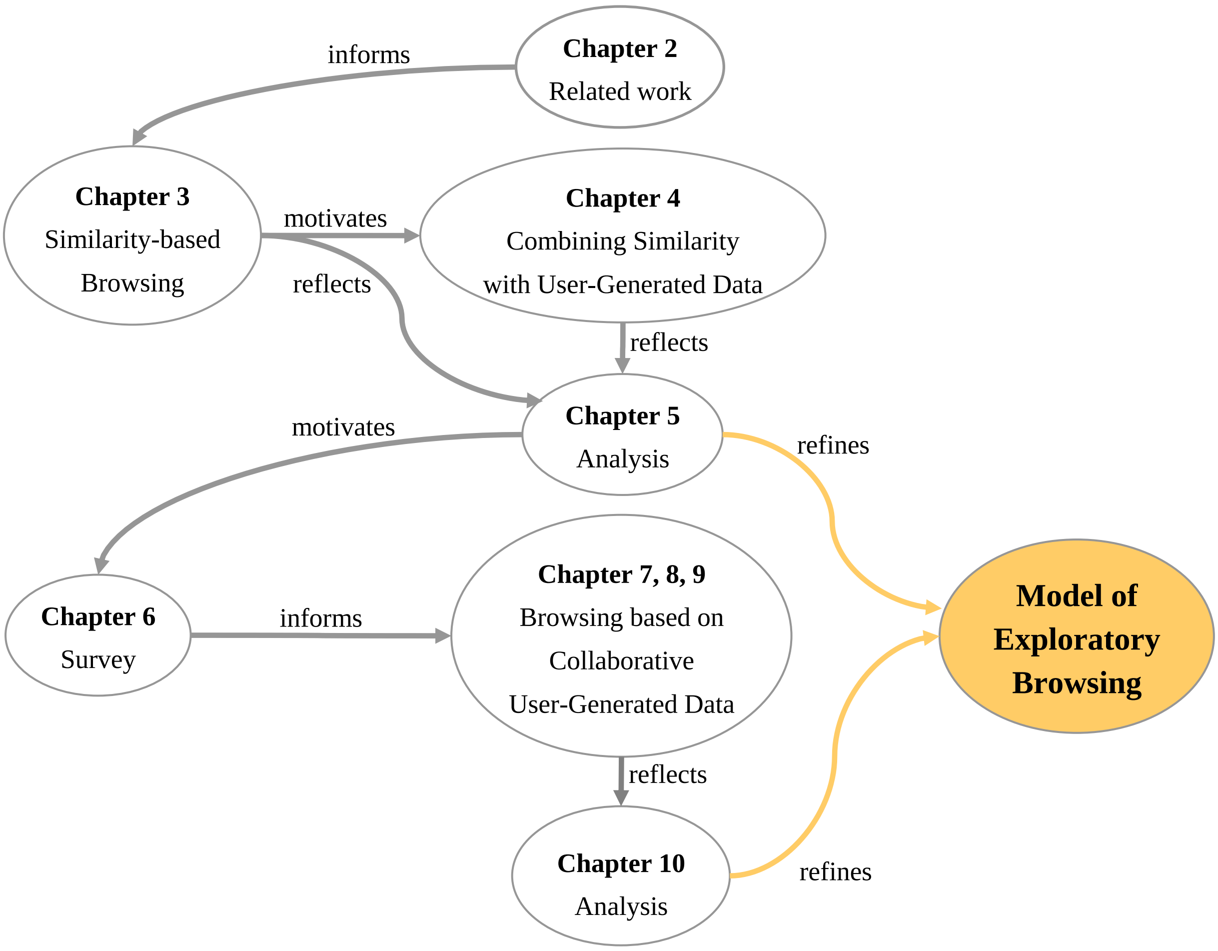


Figure 1.6: The structure of this thesis.

This thesis consists of four parts and the structure presented in this section (also shown in Figure 1.6) illustrates the exploration process of this thesis:

Part I forms the foundation of this research.

Chapter 1: Introduces the motivation, research context and methodology of this thesis. The definition and initial model of *Exploratory Browsing* are also presented.

Chapter 2: Offers a review on related work in the field of interface design for media collections. A brief overview of people' behavior around photo and music collections is given respectively, followed by a discussion of representative applications.

Part II presents our investigation to support the *Exploratory Browsing* experience with personal media collections.

Chapter 3: Illustrates our initial exploration based on similarity. Two prototypes are introduced: *PhotoSim* (Chen and Butz, 2008) and *MusicSim* (Chen and Butz, 2009) for similarity-based browsing in photo and music collections respectively.

Chapter 4: Presents our progress of combining similarity with User-Generated Data (UGD). Two prototypes are presented: *CloudMonster* (Chen et al., 2009a) and *PhotoMagnets* (see Figure 1.5c) for enhancing browsing flexibility with music and photo collections respectively.

Chapter 5: Discusses the implications for interface design and evaluation methodology, which are derived in the development and evaluation of the four prototypes. The model of *Exploratory Browsing* is refined based on a discovery of the transitional relations between *General Purpose Browsing* and the other two browsing activities.

Part III shifts the usage context from personal collections to online communities.

Chapter 6: Reports a survey of user behavior and relevant collaborative UGD in such an environment.

Chapter 7, 8 and 9: Present four prototypes built based on collaborative UGD: HisFlocks (Chen et al., 2010a) for comparing multiple users' tastes, TagClusters (Chen et al., 2009b) for improving semantic understanding of user-generated tags (see Figure 1.7a), SARA for generating personalized package recommendations (see Figure 1.7b), and MusicTrends for exploring worldwide consumption trends (see Figure 1.7c).

(a) (b) (c)

Figure 1.7: Some prototypes presented in this thesis: a) TagClusters (Chen et al., 2009b): Improving semantic understanding of user-generated tags. b) SARA: Generating personalized package recommendations. c) MusicTrends: Exploring worldwide consumption trends.

Chapter 10: Building upon the insights from the previous four chapters, this chapter discusses general considerations for interface design and evaluation approach. The model of *Exploratory Browsing* is further enriched by uncovering the transitional behavior between *Serendipitous Browsing* and the other two browsing activities.

Part IV summarizes and concludes this dissertation.

Chapter 11: Summarizes the process and methodology undertaken in the research and discusses the main contributions. It presents the model of *Exploratory Browsing*, with uncovered full range of transitional relations between the three main browsing activities (see Figure 1.8). A catalogue of general characteristics for such browsing interfaces is proposed, which is afforded as criteria to analyze the effectiveness of the presented prototypes. Several suggestions for designing interface for media collections are also proposed.

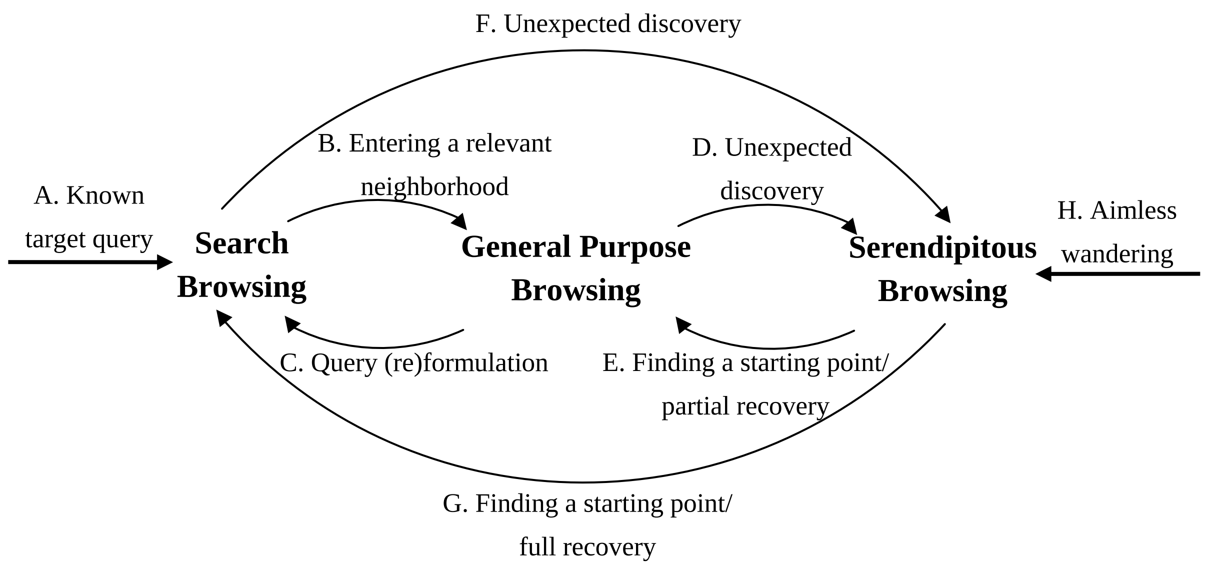


Figure 1.8: The model of Exploratory Browsing.

Chapter 12: Concludes this thesis with discussions on further work.

2

Related Work

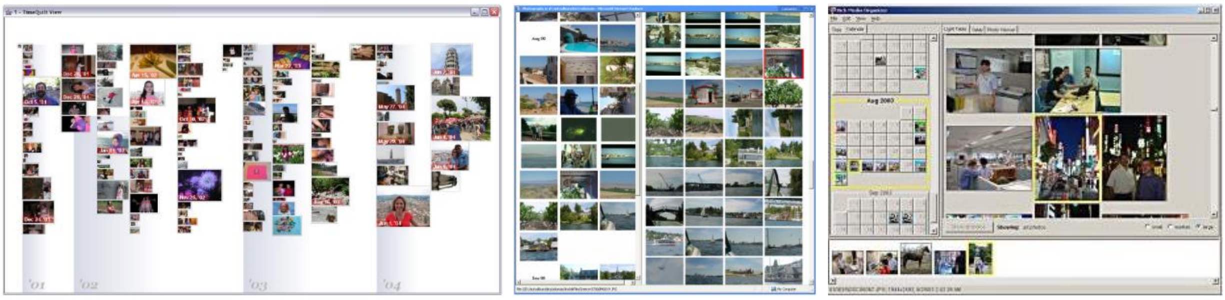
This thesis aims at designing interfaces to support the *Exploratory Browsing* experience with media collections. This chapter presents an overview of related work on interface design for media collections. We discuss photo and music collections respectively. For both media types, we explore studies on user behavior, mainly covering the aspects of organization, browsing, searching and sharing. We also investigate some relevant representative applications. Based on the review of related work, we discuss the similarity between these two media types, concerning the users' behavior and their interface needs. We also propose an initial catalogue of general interface characteristics to support *Exploratory Browsing*. We utilize this catalogue as criteria to analyze the effectiveness of some representative applications. This analysis reveals an absence of interfaces to effectively support the *Exploratory Browsing* experience with media collections. It also embodied the necessity to systematically explore the full range of browsing activities. The discussion presented in this chapter serves as fundament for our exploration in this research.

2.1 Photos

We categorize the main activities around photo collections as organization, browsing, search, tagging and sharing. For each category, we discuss the users' behavior and relevant representative applications.

2.1.1 Organization

Frohlich et al. (2002) compared people's behavior with analogue and digital photos. They discovered that most people prefer their photo collections to be organized automatically along available metadata, such as folder hierarchy or capture time, which are the main principles for organizing photo collections (Rodden and Wood, 2003; van House et al., 2004; Kirk et al., 2006). People tend to keep digital photos organized by a “roll”, which includes photos taken around the same event (Rodden and Wood, 2003). Some works allow automatic event detection and segmentation, with one selected representative photo for each event. Time Quilt (Huynh et al., 2006), Phototoc (Platt et al., 2002) and FXPAL[5] are three representative examples (see Figure 2.1).



(a) (b) (c)

Figure 2.1: Event-based organization of photo collections: a) Time Quilt (Huynh et al., 2006) breaks the time line into columns, in which the event-representative photos are displayed vertically by a space filling strategy. b) In Phototoc (Platt et al., 2002), by choosing one event representative photo in the left panel, all photos in the same event are shown in the right panel. c) FXPAL[5] offers multiple organization options, such as chronologically in a calendar view, or along different categories of events, people, places and labels in a tree view.

To illustrate the user’s life experience, researchers seek for richer representation of photo collections. Content analysis has become one of the new features and some commercial products have already included image processing to provide additional criteria for organization. Aperture Aperture[6], for example, can automatically group images with similar content into one stack, thereby identify series of motives. Besides ordering photos chronologically, it has recently become a fashion to organize photos geographically. Displaying photos by their locations for example on a map is intuitive and easy to understand. Photo Explorer (Snavely et al., 2006) creates a novel 3D representation of photos based on an automatic computation of each photo’s viewpoint (see Figure 2.2a). PhotoField (Fujita and Arikawa, 2007) allows the user to create an animation on a map, which has a visual effect of moving from one photo to another (see Figure 2.2b).

[5] http://www.fxpal.com/?p=PhotoApplication

[6] http://www.apple.com/aperture

Besides manually creating a map-based representation such as in PhotoField, some applications, such as the commercial products iPhoto[7] and Picasa[8], automatically display photos on a map based on their geo-coordinates. This geographic information is universal and bears good reproducibility. However, they cannot efficiently narrow the semantic gap between the physical coordinates and the high-level perceptions (Toyama et al., 2003). To enhance the effectiveness of location-based representation, Toyama et al. (2003) combined time and location information in WWMX (see Figure 2.2c). To improve the understandability of geo-coordinates, Naaman et al. (2004b) spent much effort in naming and mapping them with understandable and meaningful names. This work was later enhanced with additional contextual information such as season, light status, weather and temperature (Naaman et al., 2004a). Davis et al. (2006) combined geo-coordinates with content analysis to detect photos containing the same people. LOCALE (Naaman et al., 2003) recommends tag suggestions for unlabeled photos based on tags of other labeled photos which share similar geographic attributes.

(a) (b) (c)

Figure 2.2: Geographical organization of photo collections: a) Photo Explorer (Snavely et al., 2006) organizes photos of popular tourism sights in a 3D interface. b) PhotoField (Fujita and Arikawa, 2007) allows a manual organization of photos on a map. c) WWMX (Toyama et al., 2003) uses geo-coordinates to organize photos on a map.

2.1.2 Browsing

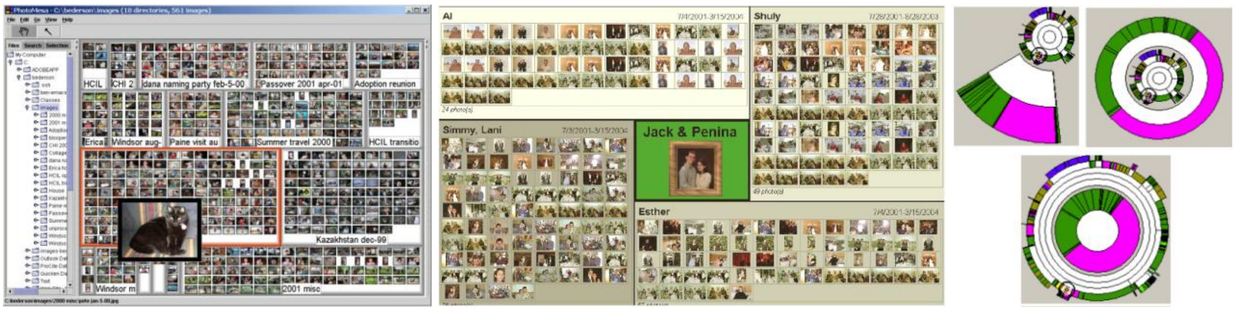
People rely heavily on time and event in search and browsing (Kirk et al., 2006). The frequency of visiting photos tends to decrease over time and people are likely to interact more often with recent items than with older ones (Rodden and Wood, 2003; Kirk et al., 2006). People rarely look for one specific item and instead, a significantly larger proportion of time is spent in browsing-like filtering rather than in explicit search activities (Drucker et al., 2004;

[7] http://www.apple.com/ilife/iphoto

[8] http://picasa.google.com

Kirk et al., 2006). Sometimes people might just browse casually for pleasure. During browsing, people may easily get sidetracked by serendipitous discovery. They often change their mind and they might look for a particular item but in the end select something different (Kirk et al., 2006). During the user study of a tabletop application for co-located photo sharing, Hilliges et al. (2009) observed that users were easily distracted by topically incoherent photos.

Rodden and Wood (2003) suggested that one of the basic functionalities for photo management software is to offer a good overview for large collections. Challenged by the enormously increasing sizes, researchers try to create novel interfaces to support efficient browsing among large amounts of media items. One strategy is to summarize the collections with representative photos (Huynh et al., 2006; Platt et al., 2002). Another option is to maximize the usage of the screen, present as many images as possible in one time, and offer details by zooming in. PhotoMesa (Bederson, 2001) is such a Zoomable User Interface (ZUI). It adopts a space-filling strategy based on a quantum treemap (Shneiderman, 1992). Each directory is displayed in a different size, determined by the number of photos inside (see Figure 2.3a). Combs and Benderson (1999) experimented with 2D, ZUI and 3D interfaces. They found a comparable performance with small numbers of photos, and that 2D and ZUI yielded lower selection errors with large numbers of photos. Kustanowitz and Shneiderman (2005) proposed a radial representation by placing a primary region in the center and groups of relevant photos around in a radial fashion (see Figure 2.3b). Stasko and Zhang (2000) developed a circular layout to represent the hierarchical structure of a collection. To enhance the readability in depth, they proposed three visualization concepts with different placements of focus and context (see Figure 2.3c).

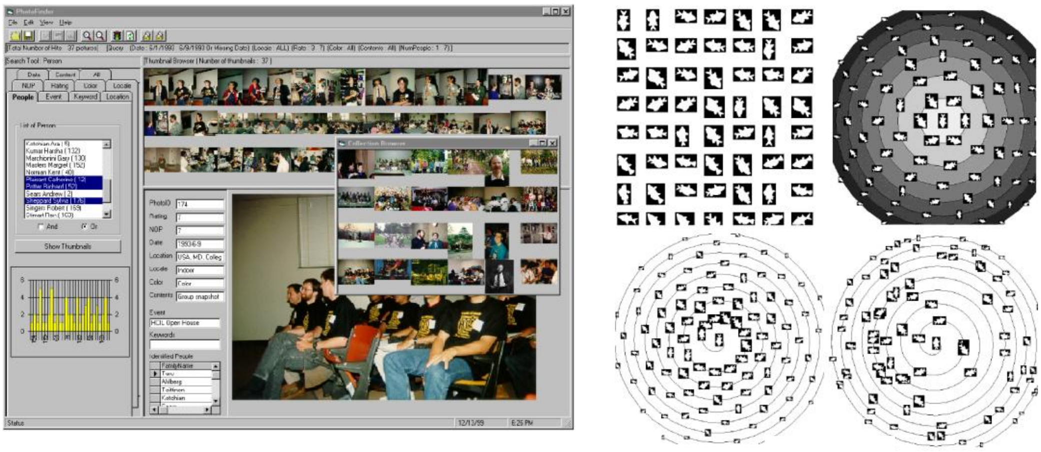


(a) (b) (c)

Figure 2.3: Representation of large photo collections: a) PhotoMesa (Bederson, 2001). b) Radial representation (Kustanowitz and Shneiderman, 2005). c) Three visualizations for focus and context (Stasko and Zhang, 2000): The context is shrunk with expanded focus. Top left: the context is pushed to the opposite side from where the focus region grows. Top right: the context is placed in the center, and the focus is displayed as a ring around it. Bottom: the focus is shown in the center and the context is pushed outward.

2.1.3 Search

People frequently look for photos from one event or from different events but sharing common attributes (Rodden and Wood, 2003). Thereby, actual search space is relatively limited in a few candidate folders (Kirk et al., 2006). Exploratory behavior is very common with media collections. Finding a target in a directed search is not a typical activity around photo collections (Kirk et al., 2006). Therefore, it is inappropriate to assess effectiveness of such interfaces by speed. People are satisficing rather than optimizing their selection (Bentley et al., 2006). They attempt to find an acceptable alternative rather than the optimal solution.



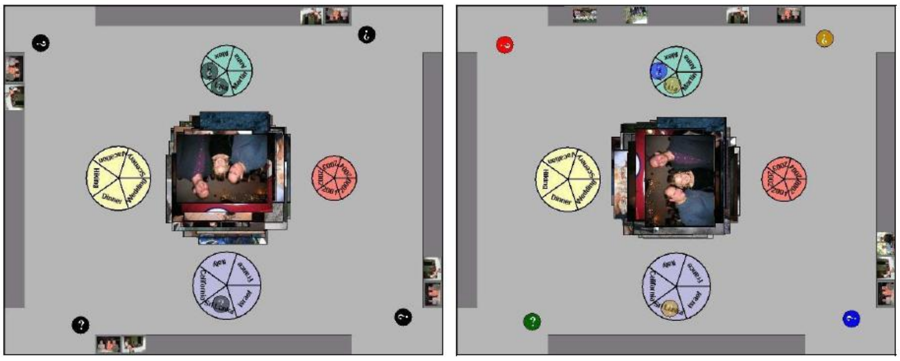
(a) (b)

Figure 2.4: Search in photo collections: a) PhotoFinder (Kang and Shneiderman, 2000). b) Three representations of retrieval results (Torres et al., 2004): grid representation (Top left), concentric rings (Top right), spiral placement based on image ranking (bottom left) and similarity (bottom right).

PhotoFinder (Kang and Shneiderman, 2000) allows searching for photos with Boolean query. Multiple search criteria are offered, such as people, time, location and color (see Figure 2.4a). The main problem with query-based search is that photo collections initially lack meaningful names and people are reluctant to annotate their photos manually (Rodden and Wood, 2003). Some researchers aim to enhance search efficiency with content analysis, such as example-based query, searching for photos that are visually similar to an example photo (Boll et al., 2007). MediAssist (O'Hare et al., 2006) employs content analysis to suggest labels of people and building/non-building. The still imperfect content analysis techniques may hinder the practical employment of low-level features for search (Markkula and Sormunen, 2000). To enhance the understandability of retrieval results, Torres et al. (2004) proposed two visualizations of spiral and concentric rings (see Figure 2.4b). For a query based on multiple criteria, MediaGLOW (Girgensohn et al., 2010) groups the retrieval results into different clusters to enable a "multi-faceted" navigation within the results. Rodden et al. (2001)

compared the effectiveness of different spatial layouts and suggested that organizing photos by category was more understandable than global visual features, such as color and textures. A similar conclusion was presented by Janiszewski (1998).

Concerning the tools supporting co-located photo search, most of them are tabletop applications. Morris et al. (2006a) presented TeamSearch to support co-present collaborative search (see Figure 2.5a). Two modes of query are provided: Integrating all group members' needs into a single query (collective mode), or interpreting them as individual requests (parallel mode). The evaluation results indicated a comparable efficiency of both strategies, and that the collective mode had an advantage in enhancing group collaboration and awareness. Shen et al. (2001, 2002) developed Personal Digital Historian (see Figure 2.5b) for organizing digital collections based on 4W categories (who, where, when, what). User can select items in different categories and combine them for filtering. The user can get relevant photos for each photo, for example, those containing the same person or sharing the same keywords.



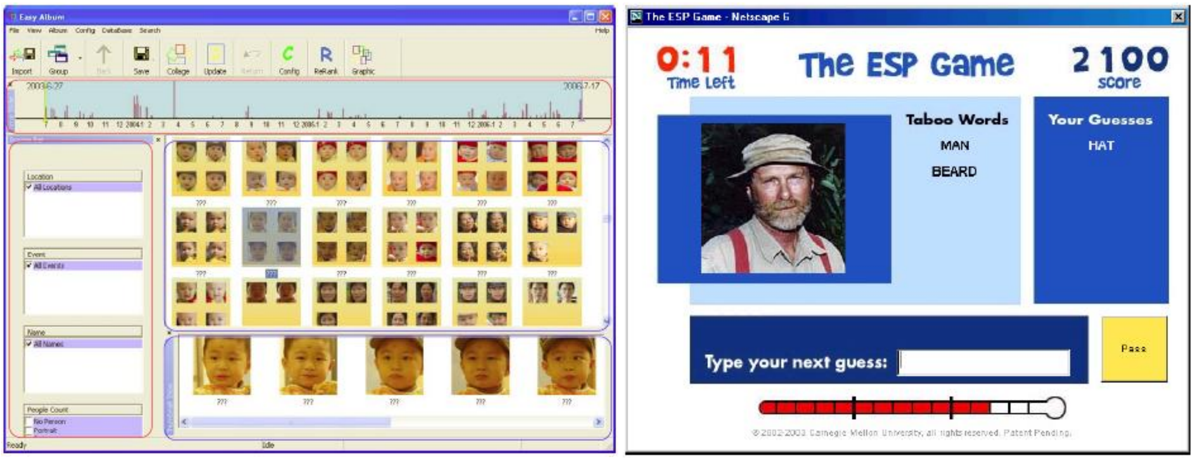
(a)

(b)

Figure 2.5: Co-located search: a) TeamSearch (Morris et al., 2006a) provides two search modes: Collective mode (left) and parallel mode (right). b) Personal Digital Historian (Shen et al., 2002) offers different views in a circular interface: who (left), where (middle) and when (right).

2.1.4 Tagging

Based on an analysis of different audiences, Kustanowitz and Shneiderman (2004) grouped the tagging motivations into four categories: self, family and friends, colleagues and neighbors, citizens and markets. According to the distribution of efforts between the user and the system, tagging techniques can be grouped into three types: manual, semi-automatic and automatic. FotoFile (Kuchinsky et al., 1999) allows the user to define a personal set of keywords. Kang and Shneiderman (2000) applied a drag-and-drop strategy to facilitate manual annotation of photos. EasyAlbum (Cui et al, 2007) clusters photos with similar scene together to facilitate the "bulk annotation" (see Figure 2.6a). Naaman et al. (2003, 2005) offered tag suggestions for unlabeled photos which are location- or event-relevant to other labeled photos. In MediAssist (O'Hare et al., 2006) the user can confirm or correct the results of the automatic face detection by removing false faces or adding missed ones. In Shoebox (Frohlich et al., 2002, 2004) textual tags are enriched by audio annotations. Davis et al. (2004) assisted photo annotation right after capturing, an essential moment to motivate tagging (Ames and Naaman, 2007). Von Ahn and Dabbish (2006) innovatively converted the tagging task into an enjoyable computer game (see Figure 2.6b).



(a) (b)

Figure 2.6: Applications facilitate photo tagging: a) EasyAlbum (Cui et al., 2007) offers a facility of "bulk annotation". b) Von An and Dabbish (2006) converted photo tagging into a computer game.

Since tagging is often consider a time-consuming and tedious task that people are reluctant to conduct (Rodden and Wood, 2003), researchers devote in automatic content-based labeling. iPhoto and Picasa use face recognition to identify people in the photo and automatically tag it with those people's names. Drucker et al. (2004) integrated content detectors for automatic labeling, such as similarity-, in/outdoor- and face-detectors. However, the results of such a search are not easily understood, as users often expect similarity on a semantic level. This is one of main reasons why those low-level features have not applied widely (Rodden and Wood, 2003).

Due to the free nature of tagging, there is inevitable noise and redundancy in uncontrolled tags, such as name inconsistency (Rodden and Wood, 2003). The other issue is the semantic gap between tags and high-level semantics (Davis et al. 2004; Kennedy et al., 2006). Collaborative Filtering (Goldberg et al., 1992; Mehta et al., 2007) is an effective solution to improve the quality of tags generated among a broad audience.

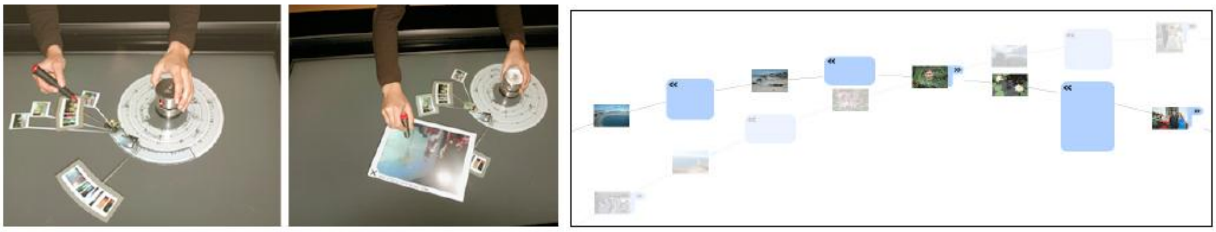
2.1.5 Sharing

Van House et al. (2004) identified photos' social functions as evoking memory, relationship maintenance and self-expression. Storytelling, communicating and sharing experience are the most common and enjoyable uses of photos (Balabanović et al., 2000; Kirk et al., 2006). People discuss and share photos frequently with their family and friends, both co-located and remotely (Miller and Edwards, 2007; Rodden and Wood, 2003).

Crabtree et al. (2004) and Frohlich et al. (2002) explored people's photo sharing behavior and discovered that co-present sharing takes place commonly in a home context with family and friends. This activity is also named reminiscing talk (Frohlich et al., 2002), a collaborative photo-talk for recalling details jointly and finding memory together. The contribution from each participant is symmetrical. Storytelling to those who were not present is a way of showing off experiences, making an impression and conveying information (Marshall and Bly, 2004). Because of the nature of one presenter and multiple listeners, the contribution of participants in this case is asymmetrical. Balabanović et al. (2000) categorized two styles of storytelling: In the photo-driven mode, "the subject explains every photo in turn, the story prompted by the existing sequence of pictures". The story-driven mode was defined as "the subject has a particular story in mind […], then gathers the appropriate photos and recounts the story". No clear distinction between these two styles was found in an observation of user behavior. Users may continuously switch between these two styles, rather than stick to one of them.

A number of applications have been developed to support collaboration and sharing of photos. Tabletop is widely applied for co-located photo storytelling and sharing. Photohelix (Hilliges et al., 2007) is a tabletop application to support collaborative sharing of photos (see Figure 2.7a). Personal Digital Historian (Shen et al., 2002) presents a circular tabletop interface to encourage conversation and storytelling about group histories (see Figure 2.5b).

Most of existing applications support co-located sharing, and few address collaboration in a distributed fashion. Remote sharing is currently limited to pass\ on photos by email, webpage or mobile phones. In PhotoArc (Ames and Manguy, 2006), the user can organize photos as linear arcs, which are connected by textual narratives. The user can also view the narratives intersecting with others, or create multiple versions of a particular narrative for different audience (see Figure 2.7b). MyLifeBits (Gemmell et al., 2005) also allows the user to create a sequence of photos by drag-and-drop. Schäfer et al. (2004) allowed users to create and share stories remotely in a shared workspace.



(a) (b)

Figure 2.7: Storytelling and sharing of photos. a) Photohelix (Hilliges et al., 2007). b) PhotoArc (Ames and Manguy, 2006).

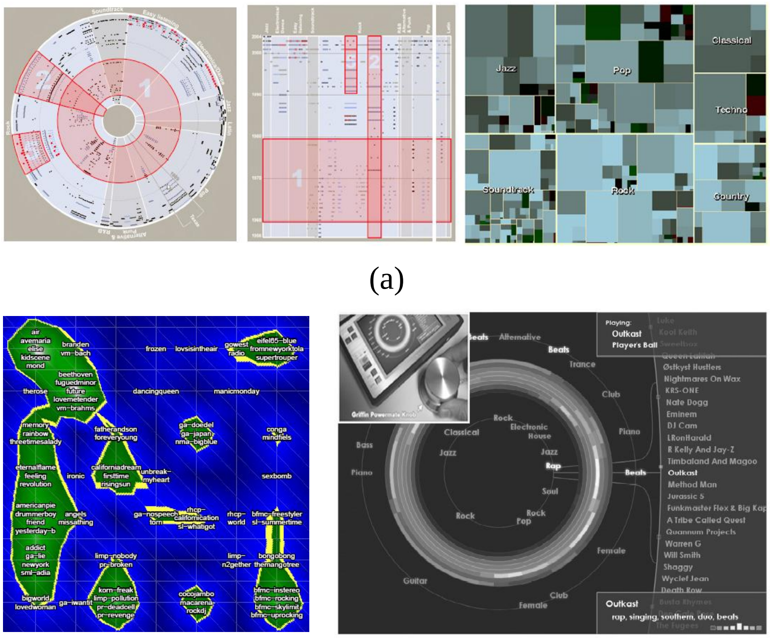
2.2 Music

We categorize the main activities around music collections as organization, information seeking, playlist generation and sharing. For each category, we discuss user behavior and relevant representative applications.

2.2.1 Organization

Vignoli (2004) investigated people's behavior with digital music collections and found that organization mainly follows the hierarchical structure of artist-album-song or genre-subgenre. Music does not fit the main organization principles may be stored in additional folders. Current commercial products display music mostly in a list, possibly visually enhanced with cover art, such as in iTunes (see Figure 1.1d). Criteria such as artist, album, track and genre, can be used for sorting or filtering. Pure list-based organization lacks an illustrative description of the relations between songs and thus cannot support flexible explorations.

Similar to photos, a good overview for a large music collection is necessary. Torrens et al. (2004) offered three visualization concepts: a disc, a rectangle and a tree-map (see Figure 2.8a). In Islands of Music (Pampalk, 2003), clusters of similar songs are represented as islands based on their low-level features (see Figure 2.8b). MusicRainbow (Pampalk and Goto, 2006) displays artists in a "rainbow" (see Figure 2.8c). Artist similarity is calculated based on the accumulated acoustic similarity of corresponding tracks. Artists are grouped into different rings by their web labels. In most of these applications, the user has no active control of the organization, and Cunningham et al. (2006) suggested that a visualization supporting rearrangement of an entire collection will be more appealing.



(a)

(b) (c)

Figure 2.8: Organization of music collection. a) Three visualization concepts proposed by Torrens et al. (2004): A disc (left), a rectangle (middle) and a treemap (right). b) Islands of Music (Pampalk, 2003). c) MusicRainbow (Pampalk and Goto, 2006) allows the user to select genres and artists by rotating a knob.

2.2.2 Information Seeking

Cunningham et al. (2003) explored the music seeking strategies in physical environment (CD shop and public library). Two main categories of activities were observed: search (locate a specific album or song) and browsing (locate CDs containing music the user might like). Bibliography is the basic information for known item search, but the user may not remember it correctly. Therefore, a facility of Query-By-Humming was suggested. Genre-artist structure helps browsing with no specific target. Unfortunately, there is no standard definition of genre and people may have different understanding of the same genres. Therefore, more intrinsic attributes should be included for browsing, such as instrumentation and rhythm, based on which Query-By-Example can be conducted. Search and browsing activities can be interleaved: A known item search can be expanded to a relatively directionless browsing, which can possibly inspire a refinement of a search. Music seekers will benefit from an interface that allows them to move seamlessly between these two seeking activities. Facilities encouraging users to discover new music should be further explored.

Vignoli (2004) categorized two main browsing activities around digital music collections: view- and association-based browsing. View-based browsing is driven by the intension of selecting a set of songs matching certain properties. Association-based browsing originates from the objective of finding songs similar to an example song.

Current browsing interface for music collection is rather simple. Cunningham et al. (2006) suggested that extensive browsing structure can help users to narrow down the choices to a reasonable sized set of candidates. Interactions will facilitate browsing and enhance the feeling of enjoyment. Beside cover art, visible metadata for music collections is limited. During browsing the user has to imagine how each song might fit the theme. Therefore, additional information such as lyrics should be easily accessible.

Boolean query is a common search method, with which users describe their musical requirements as a logic expression. However, non-expert users may have strong difficulties to express their musical preferences in a formal way (Vignoli, 2004). Frequently used attributes for search are metadata such as artist, song, album, genre, lyrics and year (Bainbridge et al., 2003; Downie and Cunningham, 2002; Vignoli, 2004). Other less common but also desired properties are mostly related to intrinsic characteristics of music, such as mood, instrumentation and acoustic similarity. They help to improve the flexibility with music selection (Brown et al., 2001a) and facilitate similarity judgments of music (Cupchik et al., 1982). Genre is one of the most common criteria for music organization and retrieval (Cunningham, 2006). It is more appropriate for classification of unknown music because it gives general information one can expect from a piece of music. For personal collections the name of artist conveys more information than genre (Vignoli, 2004).

2.2.3 Playlist Generation

As background for other activities, music should be appropriate for the context and audience. Compiling a playlist goes beyond selecting a set of individual songs. It requires skills and taste, and cannot be solved with static solutions (Lee and Downie, 2004; Downie and Cunningham, 2002). Gates et al. (2006) conducted interviews with Disk Jockeys (DJs) and found that they always create smooth musical bridges between changes, which is not a simple task, due to the ever-changing nature of the audience and the limited awareness of the individual tastes of its members.

Systems enabling automatic playlist generation focus mainly on algorithms (Pohle et al., 2005a; Logan, 2002) or the user’s listening habits (Pampalk et al., 2005; Andric and Haus, 2006). PATS (Pauws, 2002) tries to balance the coherence and variation of a playlist by assuring the same song won’t be recommended multiple times. The results of the user study showed that the playlists generated by this system outperformed those randomly assembled ones. The sequence in the playlist is crucial. Hansen and Golbeck (2009) addressed the consistency between items such as co-occurrence and similarity. Although automatic playlist generators alleviate the users’ effort, they might remove the fun of creating playlists and reduce the users’ active control over the system output (Lee and Downie, 2004). Some applications address this issue by allowing more user interference, for example, by specifying the length of the playlist (Aucouturier and Pachet, 2002). SatisFly (Pauws and van de Wijdeven, 2005) allows the user to define specific constrains, such as number of songs, genre, artist and tempo. The system applies constraint satisfaction to produce a list of songs which meet these requirements. Cunningham (2006) suggested that automatic playlist generation techniques should be relaxed to provide suggestions rather than automatic selections.

Researchers are looking for interactive graphic representations to support music selection and playlist generation. User interactions are introduced to encourage the user’s participation and enhance the quality of the final playlist. Musicream (Goto and Goto, 2005) offers a similarity-based sticking function: users can select one song, and other similar songs will be automatically attached to it, thereby forming a playlist (see Figure 2.9a). The user can adjust the order of songs in the playlist by drag-and-drop. E-Mu jukebox (Vignoli and Pauws, 2005; Herrera et al., 2005) offers five adaptors for search: tempo, sound, genre, mood and year. They are visualized in a graphics where the radial distance presents the value of weights (see Figure 2.9b). The user can change the weight by replacing the criteria. In Artist map (van Gulik and Vignoli, 2005) the user can choose any two of the four criteria of mood, genre, year and tempo to display music on a map (see Figure 2.9c). The user can draw a path and all the items in this path will be compiled into a playlist.

(a) (b) (c)

Figure 2.9: Interactive generation of playlist. a) Musicream (Goto and Goto, 2005). b) E-Mu jukebox (Vignoli and Pauws, 2005; Herrera et al., 2005). c) Arstist map (van Gulik and Vignoli, 2005).

2.2.4 Sharing

Besides personal usage, music possesses public and social functions. Through an investigation of the music sharing practices in iTunes, Voida et al. (2005) discovered that music sharing helps to create personal images before others. Brown et al. (2001b) compared music sharing behavior in offline and online sharing systems, and discovered that music sharing is tightly bonded with social activities. They suggested that music should be shared in a more collaborative and community-related environment. Brown et al. (2001a) surveyed practices with conventional and mp3 users and found that friends are one of the main sources of new music. People show positive attitude towards others' ratings, recommendations and opinions (Brown et al., 2001a, 2001b; Bentley et al., 2006). Personal familiarity with helpers is a decisive factor for seekers (Downie and Cunningham, 2002). Jacobsson et al. (2005) compared recommendations produced by friends and Push!Music, a Music Recommender System. The result indicated that friends performed generally better but sometimes the system gave fresh and unexpected recommendations, which revealed the value of having both personal and system recommendations. In MusicSun (Pampalk and Goto, 2007), artists are recommended based on example artists selected by the user. The recommendation was generated through a combination of audio-, web- and tag-based similarities.

(a)

(b)

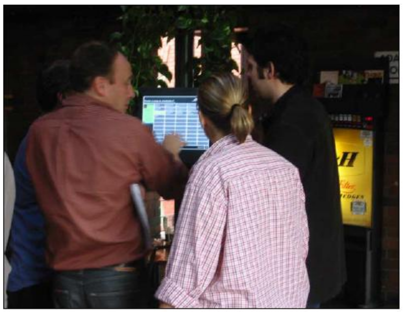


Figure 2.10: Applications support co-located music consumption. a) When users enter the space of Flytrap (Crossen et al., 2002), their musical tastes are identified through radio frequency ID badges. The system selects the song to be played next based on votes. The current playing song is placed in the center of the interface, and other songs' distances to it present their possibilities to be placed next. b) MUSICtable (Stavness et al., 2005) allows users to collaborative decide on the portion of music to be played on a map. The eight buttons around the table allow users to collectively control the browsing direction. c) With Jukola (O'Hara et al., 2004), people can vote for the next song through handheld devices or a public display.

Music conveys cultural information, provokes emotions, and enhances the experience of visiting a social place (Komninos et al., 2008). In many social settings employing a DJ is expensive, and music selection becomes the responsibility of either amateur volunteers or, in most cases, relies on pre-determined playlists which are left running on their own. The latter static solution makes music inflexible and disjointed with the context in a social environment. In most existing applications for a co-present environment, such as in musicFX (McCarthy and Anagnost, 1998), Bluemusic (Mahato et al., 2008) and Flytrap (Crossen et al., 2002, see Figure 2.10a), user input is restricted to pre-defining personal preference. In other applications, users can dynamically interfere with the system recommendations, for example, by changing their genre preference in AmbientDJ (Komninos et al., 2008) or dynamically nominating songs in Jukola (O'Hara et al., 2004, see Figure 2.10c). In most of those systems, playlists are generated based on votes. MUSICtable (Stavness et al., 2005) is a tabletop application which allows collaborative music selection on a map (see Figure 2.10b). Some applications, such as BluetunA (Baumann et al., 2007) and Push!Music (Jacobsson et al., 2005), allow sharing music interests with people nearby.

2.3 Discussion

This chapter has presented a review on user behavior and relevant representative applications. It formulates the basis of our understanding of the users' fundamental needs and state of the art in the area of designing interfaces for media collections. This section discusses the behavioral similarity between photos and music. An analysis of the representative systems reveals a lack of interfaces to effectively support the *Exploratory Browsing* experience with media collections. It also shows the necessity to systematically explore the full range of browsing activities.

2.3.1 Behavioral Similarity with Media Collections

Photos and music are two different media types, with which people have different behavior and thus forming different interface requirements. Although most of the existing work deals with photos and music respectively, we notice in our survey that these two media types share similarity in the aspects of user behavior and interface requirements.

Organization: The common organizations of photos and music both follow the most reliable principles, such as time for photos and artist-album-song for music. For a rather large collection, a good overview is helpful to obtain an overall impression of the entire collection and to locate possible areas of interest.

Social function: Both photos and music convey social information. They help to share experience, maintenance personal image before others, and reinforce social connections with family and friends. People like to share and tell stories with their media collections, both co-located and remotely. Most of existing applications supporting collaborations focus on a co-located environment, and few explore the context of remote sharing.

Information seeking: There are two main activities with media collections: known item search and browsing. People have difficulties to formulate their requirements formally. There is often a semantic gap between query terms and high-level perceptions. Search for one specific item is not a typical activity with media collections. Instead, people tend to spend more time on browsing, since it is easier to recognize one qualified item on the way of browsing than construct a specific query. Instead of look for the optimal solution, the selection strategy is rather flexible and people tend to make a choice among several candidates. During browsing, people often change their mind and easily get sidetracked by unexpected discoveries.

2.3.2 Interface Requirements to Support Exploratory Browsing

Based on the discussion of supporting *Exploratory Browsing* with media collection (cf. Section 1.2.3 in Chapter 1), we believe that such interfaces should follow several fundamental characteristics. As a specific target is involved in *Search Browsing*, keyword-based search should be supported. For collections initially lacking meaningful names, such as photos, the system should allow a structured navigation, which helps the user locating the target quickly along a clear structure. Search for one specific item is not common with media collections, and people tend to spend more time on browsing. They may start with a relevant neighborhood and invest effort on comparing and understanding the underlying relations between relevant items, which helps users to find acceptable candidates. Therefore, allowing definition of areas of interest is necessary. Since people's mental model of their collections often goes beyond a linear structure, the system should provide additional nonlinear representations of the collection. Moreover, these representations should be adjustable, in order to follow people's often changing intentions. Without specific goals, people may easily get sidetracked by unexpected but interesting findings. To help them recovering from serendipity, backtracking should be allowed.

Based on these general interface characteristics, we analyzed the effectiveness of some existing work to support *Exploratory Browsing*. Most of the photo applications shown in Figure 2.11 support structured navigation, mainly based on time or folder structure. Regarding the nonlinear representation, several systems provide a map-based visualization based on photos' geographic information or content similarity.

Linear structures for music collections, such as artist- and album-views in iTunes, can be easily achieved. In Figure 2.12, one can observe that researchers are seeking for novel representations for music collections, such as a map view based on additional criteria of similarity, tempo or mood. Torrens et al. (2004) proposed three novel concepts to visualize a music collection: a disc, a rectangle and a treemap. In Artist map the user can decide on the two criteria used to display songs on a 2D map, but has no influence on the layout. Moreover, most of the applications use similarity as the only criterion for organization and browsing.

Application / Characteristic	Event-based		Geographic			Tabletop		Others	
	Time Quilt	FX PAL	Photo Explorer	Photo Field	WW MX	PDH	Photo helix	iPhoto	Media GLOW
Search Browsing									
Structured navigation	√	√		√	√	√	√	√	
Exploratory Browsing									
Nonlinear representation			√		√	√		√	√
Multiple representations		√			√	√		√	
Defining areas of interest									√
Steerable representation		√							
Backtracking									

Figure 2.11: An analysis of some photo applications based on the initial catalogue of general interfaces characteristics to support Exploratory Browsing.

Application / Characteristic	Similarity-based		Playlist generation			Others			
	Islands of Music	Music Rainbow	Music-ream	E-Mu jukebox	Artist map	iTunes	Torrens	Music Land	MUSIC table
Search Browsing									
Structured navigation						√			
Exploratory Browsing									
Nonlinear representation	√	√		√	√		√	√	√
Multiple representations							√		
Defining areas of interest				√					
Steerable representation				√					
Backtracking			√						

Figure 2.12: An analysis of some music applications based on the initial catalogue of general interfaces characteristics to support Exploratory Browsing.

In most of the systems listed in Figure 2.11 and 2.12, the user cannot actively control the representation nor define areas of interest. The three exceptions are FXPAL, MediaGLOW and E-Mu jukebox. In FXPAL the user can refine the automatically generated event structure. MediaGLOW allows defining a time interval of interest. E-Mu jukebox allows defining music preferences by drag-and-drop the relevant parameters in a graph, and the matching songs will be dynamically updated in a list. To facilitate recovery from serendipity, Musicream is the only system to allow recalling a past activity with its “time machine” facility. It records all operations performed by the user as well as all screen changes. The user can switch back to a previous status with a slider, or search for past screens where a specific song or songs similar to an example song were listened to.

In summary, structured navigation can be easily achieved based on the original reliable metadata, such as time for photos, and artist and album for music. Researchers have realized the usefulness of nonlinear exploration, but mainly achieve map-based representations based on additional criteria associated with media intrinsic attributes, such as similarity, tempo and mood. This type of additional information enables more browsing possibilities, but novel representations and more elaborate user interactions need further investigation. Other functionalities to facilitate *Exploratory Browsing*, such as defining areas of interests, steerable representation and backtracking facility, are poorly support in these systems.

Although existing studies on user behavior have already highlighted the prevalence of exploratory behavior around media collections, applications efficiently supporting *Exploratory Browsing* are still in absence. The simple and fixed browsing structures offered in current interfaces cannot efficiently support users' exploratory and often unstructured browsing activities. Moreover, fluid and elaborate user interactions are poorly supported. The full range of browsing activities and their underlying relations have not been thoroughly studied. All these facts motivate us to explore interface design to enhance the browsing experience with media collections. In this thesis we systematically examine the model of *Exploratory Browsing* in two main usage contexts: with personal collections and in online communities. The insights derived in the development and evaluation of our prototypes enable reflections and refinement of the initial model of *Exploratory Browsing*. Next part presents our exploration to support the *Exploratory Browsing* experience with personal media collections.

Part II

Supporting Exploratory Browsing in Personal Media Collections

The first part of this thesis analyzes the related work in the field of interface design for media collections, which embodies the necessity to design interfaces to support *Exploratory Browsing* in media collections. This part presents our exploration with personal media collections.

Chapter 3 discusses our initial work based on similarity and presents two similarity-based interfaces *PhotoSim* and *MusicSim* for photos and music respectively. The results of the preliminary tests revealed that similarity is appreciated as a secondary criterion for organization and browsing and should be offered on demand. We suggested enhancing the performance of similarity by a close coordination with reliable criteria and by introducing more elaborate user interactions.

To improve the browsing efficiency and flexibility, in Chapter 4 we combined similarity with reliable criteria and additional information, specifically User-Generated Data (UGD). UGD used in this chapter is generated by single users, such as item popularity extracted from the user's consumption records and tags contributed by the user. We developed two prototypes *CloudMonster* and *PhotoMagnets* to enhance the browsing flexibility with personal photo and music collections respectively, in which elaborate interactions are proposed to enable more active user control.

Chapter 5 discusses the insights gathered during the development and evaluation of these four prototypes. We proposed several general considerations for designing browsing interfaces for media collections. We also discussed the different evaluations approaches of these prototypes. The analysis and observation of the users' behavior enabled a refinement of the initial model of *Exploratory Browsing* by uncovering the transitional relations between the prominent *General Purpose Browsing* and the other two browsing activities.

The exploration in this part motivated an investigation in the other usage context in online communities, which is discussed in Part III.

3

Similarity-based Browsing

As discussed in Chapter 1, current browsing interfaces for media collections mostly offer linear scan corresponding to the underlying technical concepts (folder, events, etc.) or the metaphors borrowed from the analogue counterparts (e.g., stacks, film rolls). However, people's mental model of their media collections often goes beyond this linear structure. For example, we might look for a nice photo of our kids without caring where and when it was taken, or compile a happy and lighthearted playlist for a birthday party. This kind of similarity-driven exploration is poorly supported in today's interfaces for media collections. We started our initial exploration based on similarity. We employed similarity as the main criterion and developed *PhotoSim* (Chen and Butz, 2008) and *MusicSim* (Chen and Butz, 2009) to support similarity-based browsing with personal photo and music collections respectively. The results of the preliminary tests revealed that similarity is generally appreciated but more suitable as an additional criterion offered on demand.

3.1 PhotoSim: Similarity-based Photo Browsing

The traditional interface for photo management (e.g., iPhoto and Windows Explorer) uses a two-panel strategy. When the user selects a specific folder in one panel, all corresponding photos are shown as thumbnails in the other panel. More elaborate approaches include treemaps and zoomable interfaces, such as PhotoMesa (Bederson, 2001). Content analysis has already been integrated in some commercial products. For example, Aperture automatically stacks successive pictures which share a similar color histogram. This may speed up browsing and searching by partitioning the entire set into reasonable subsets.

Our goal was to go beyond this simple pre-grouping mechanism. We aimed to create an interface which employs content analysis to further structure photo collections below the level of time, and also allows manual overrides of this automatic organization. Since folder and time are the main principles for organization and browsing (Rodden and Wood, 2003; Kirk et al., 2006), we provide two coordinated views, one of which corresponds to the classical way of structuring photo collections along folder and time. The other one uses similarity as the main criterion.

3.1.1 Interface

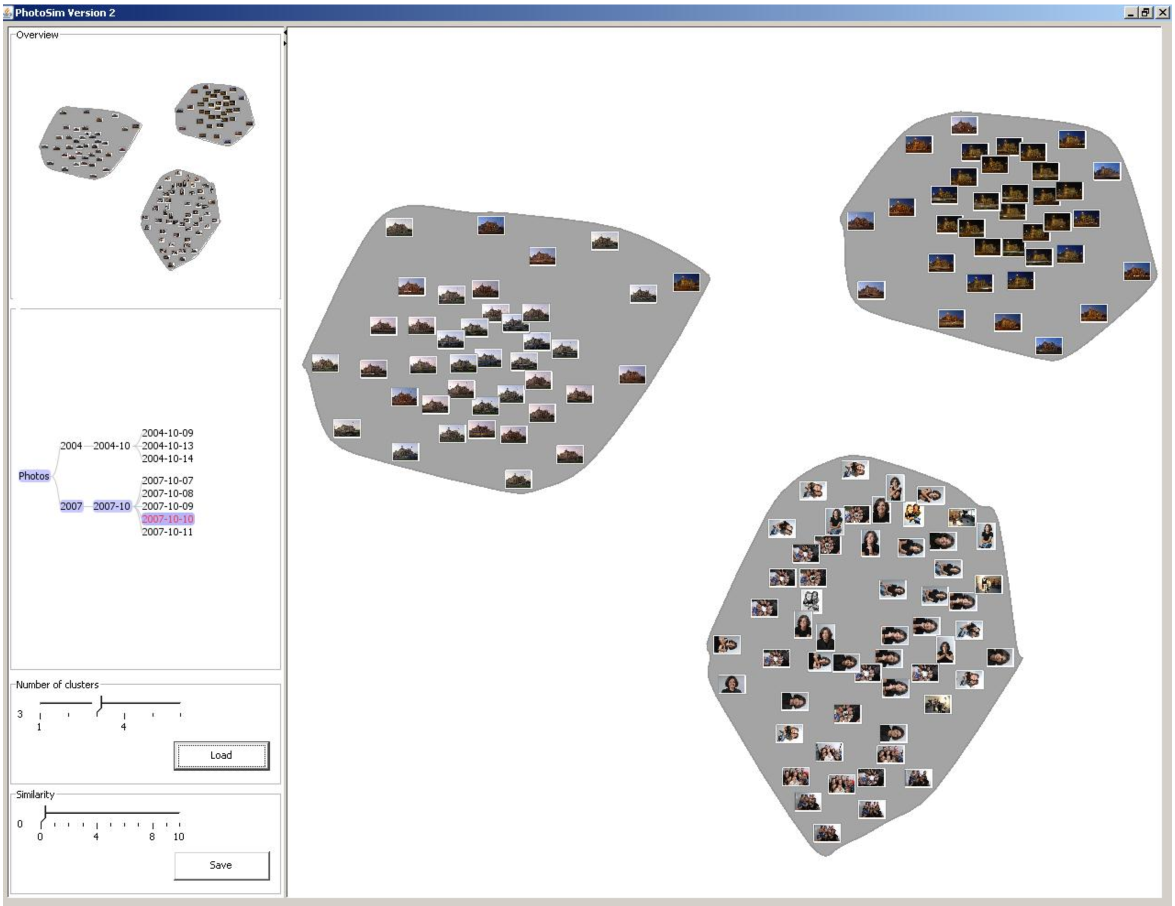
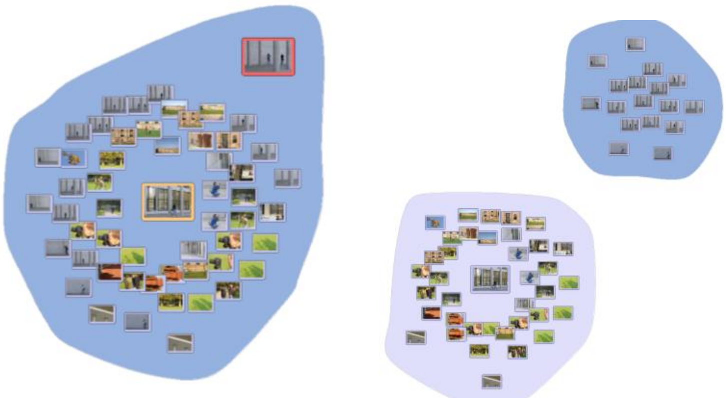


Figure 3.1: The interface of PhotoSim contains a main view on the right and an overview view, a tree view and two control panels on the left.

PhotoSim is a prototype for browsing photo collections based on similarity (see Figure 3.1). It is implemented based on the Prefuse toolkit (Heer et al., 2005) for interactive Information Visualization. The interface contains five panels: a main view, an overview, a tree view and two control panels. When a day, month, year or the entire collection is selected in the tree view, the corresponding photos will be shown in the main view, and grouped into different clusters based on their visual similarity. Photos can be clustered in this way on different levels of the temporal hierarchy. This allows, for example, finding similar photos across different days, months or even years. The user can pan and zoom freely within the main view, and is always provided with an overview in the top left corner, which can be used to jump directly to one specific cluster.

Adjusting the number of clusters or the similarity threshold in the control panels can trigger an automatic reorganization of photos in the main view. The positions of photos or clusters can be changed by drag-and-drop. The user can merge two clusters by dragging one into the other. One photo can be dragged outside its original cluster and a new cluster will be automatically created which contains this example photo and other similar ones from the original cluster (see Figure 3.2). By setting the similarity threshold to zero, the user can manually create new clusters. The user can select the “undo” menu to cancel these modifications.



(a) (b)

Figure 3.2: Example-based cluster splitting (Chen and Butz, 2008): a) A photo (highlighted with red frame) is dragged outside its original cluster. b) A new cluster is automatically created which contains this example photo and other similar ones from the original cluster.

Underlying Image Analysis

Color and texture features are used to calculate similarity between photos. 48 color features are computed in YUV color space. Since we are more interested in the color similarity than the brightness of the image, the luminance component (Y) is ignored. The two chrominance components U and V are divided into 6 sections, which lead to a color histogram with 36 dimensions. Four statistical features are extracted for each section of U and V components (Stricker and Orengo, 1995): mean, standard deviation, skewness and kurtosis. Besides color-relevant information, Haralick texture features are extracted (Haralick et al., 1973). Four distance 1 gray-tone spatial dependence matrices P_H, P_V, P_{LD} and P_{RD} are first generated, which represent the spatial relations of gray tones in the four directions of horizontal, vertical, left and right diagonal respectively. Based on these four matrices, 14 textural features are extracted for each direction. Some of them represent textural characteristics of the image, such as homogeneity, contrast and correlation. Others illustrate the complexity and nature of gray-tone transitions in the image. In total, 104 features are extracted. Due to the computational complexity, feature extraction did not work in real time. The preprocessing of a collection of 813 photos (2M Pixel each) took about 2 hours.

After extracting these low-level features, photos are grouped into different clusters based on the Simple K-Means clustering algorithm (Kanungo et al., 2000). Figure 3.1 shows the clustering results of 3 motives: portrait, building in daylight and at night. Depending on the actual degree of similarity between and within photo motives, this approach might still create too many or too few clusters in the general case. Therefore user feedback is essential to enhance the performance of such automatic solutions.

3.1.2 Preliminary Evaluation

We tested *PhotoSim* with five photo collections. Four of them were taken by four amateur photographers in a three-day photography workshop. The size of these collections varied between 317 and 812, with an average size of 507 photos. Another collection was offered by a professional photographer, containing 603 photos which were captured over five years.

Although we did not conduct a formal user study, the discussions with the authors of our example photo collections were quite encouraging. The feedback on similarity-based browsing was overall positive. The professional photographer said that he would like to use such a photo browser even more, if the structure of the collection after clustering and manual modifications could be translated back into the directory structure in the file system. He claimed that this might be a very valuable tool for temporally reorganizing his entire photo collection beyond the pure chronological order. The participants noticed that the performance of clustering algorithm was unstable and largely influenced by the content difference among motives, which needs further improvement. Currently, similarity is employed as the only criteria below the level of time. The participants desired a closer coordination between similarity and time, especially when a large time span was selected.

3.2 MusicSim: Similarity-based Music Browsing

The positive feedback towards similarity in *PhotoSim* encouraged us to apply the same concept to music. We developed *MusicSim* to support similarity-based browsing with music collections. One general problem of many existing approaches based on content analysis is that some items may be wrongly clustered, which is hard to avoid with fully automatic algorithm. Another issue is that the user can only browse passively, but has no active control over the layout. Similar to *PhotoSim*, in order to utilize human perception to improve the system's performance, user feedback is enabled through interactions in *MusicSim*. We employed audio analysis to automatically structure music collections based on similarity, but allow manual overrides of this organization by inferring feedback from the user's interaction.

Additional information, such as genre and album art, is integrated to enhance the non-visual music perception.

3.2.1 Interface

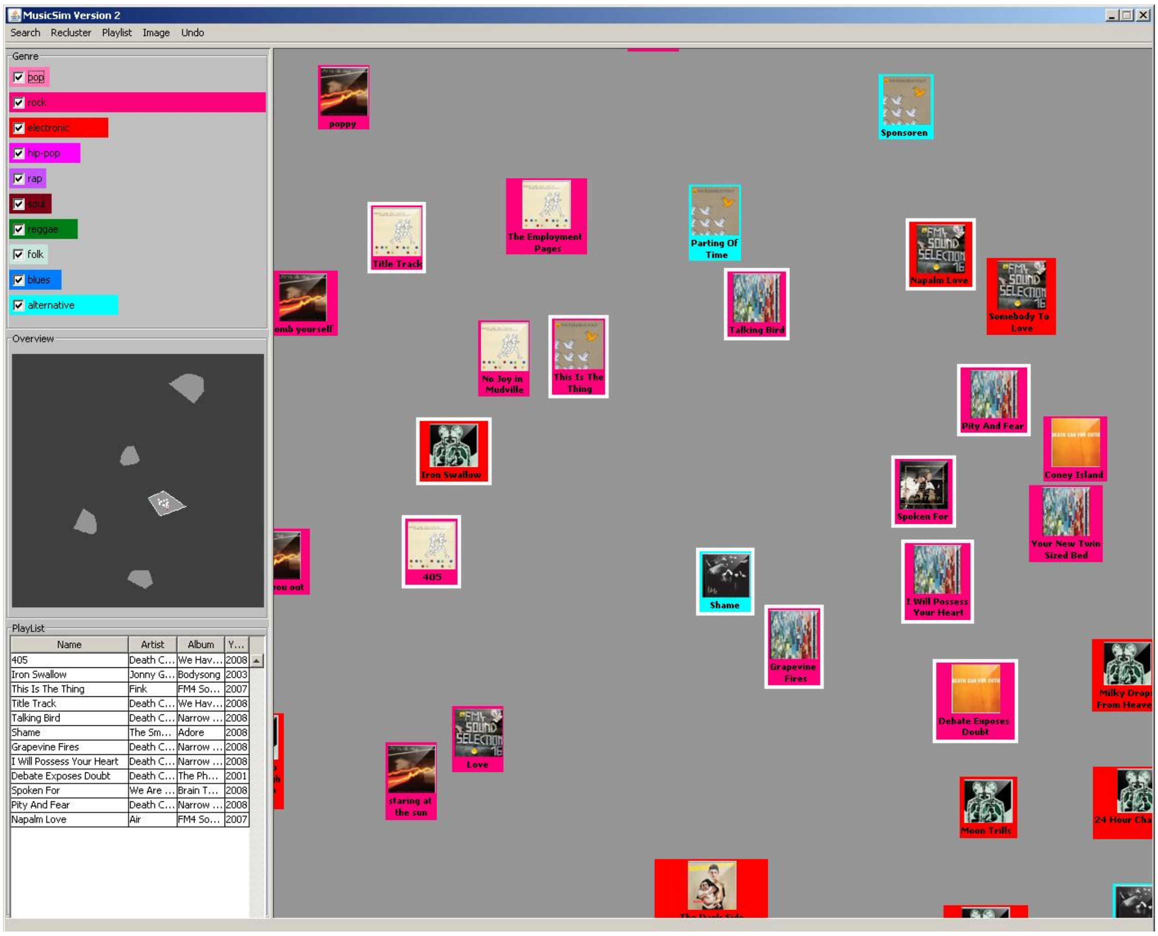


Figure 3.3: The interface of MusicSim (Chen and Butz, 2009) contains a main view on the right, and a genre histogram, an overview and a playlist panel on the left.

MusicSim is implemented based on the Prefuse toolkit (Heer et al., 2005). The interface contains four main panels (see Figure 3.3): a main view, a genre histogram, an overview and a playlist panel. Acoustically similar songs are grouped into clusters and their locations are determined by their similarity. The user can conduct keyword-based search to locate specific songs or jump to a specific cluster in the overview. The user can select multiple songs and generate a playlist for them through a shortcut menu triggered by the right mouse button.

Visual Assistance for Music Perception

In contrast to photos, music itself carries no visual information, and the user cannot describe what a song sounds like without listening to it. We consider the enhancement of music perception in several ways. Since cover art helps visually reminding of an album (Kim and Belkin, 2002; Vignoli, 2004), it is displayed for each song, and more detailed information can be shown with mouse hovering, such as artist, album and year. Genre is one of the most common criteria for music organization and retrieval, which conveys general information about what the user can expect from a specific piece of music (Cunningham, 2006). Therefore, in *MusicSim* each song is color coded with its genre. Although there is no commonly accepted mapping between genres and colors, we try to match at least some obvious pairs or use lexical similarity for the color assignment, such as red for rock and blue for blues. Beyond this, we tried to symbolize genres with drastic emotion by brighter colors and peaceful ones by darker colors. Since musical perception is very subjective, this matching is still bound to remain partly arbitrary for different users, and it can be further personalized for example by allowing the user to assign the colors. A genre histogram of the entire collection is provided, which can be used for genre filtering.

Integration of User Feedback

Music understanding is complex and subjective. It will never be possible to perfectly match a user's musical perception by a fully automatic approach. One solution is to take the human into the loop and combine audio analysis with user feedback. Similar to *PhotoSim*, we provide manual overrides of all automatic mechanisms. The simplest override is to drag a song from its original cluster to another cluster, or to place it in the blank area to create a new cluster. On top of this, there are two ways to reorganize clusters. For the manual reorganization, the user can split one existing cluster into two sub-clusters by drawing a line across it, or merge two clusters by dragging one into the other. For the automatic reorganization, the user can adjust the number of clusters, and all songs will be re-clustered accordingly, resulting in a new layout. All these modifications can be withdrawn through an "undo" functionality.

Underlying Audio Analysis

A set of low-level audio features is extracted for each song, using jAudio (McEnnis et al., 2005), an audio feature extraction toolkit. Taking an evaluation of the frequently used audio features (Pohle et al., 2005b) as reference, we extracted 11 main categories of low-level features: spectral centroid, spectral rolloff, spectral flux, compactness, Root Mean Square (RMS), fraction of low amplitude frames, zero crossings, Mel-Frequency Cepstral Coefficients (MFCCs), Linear Predictive Coding (LPC), beat histogram and method of moments.

Spectral centroid represents the center of the magnitude distribution of the spectrum. Spectral rolloff is the frequency under which 85% of the energy in the spectrum falls. It is an indicator of the skewness of the frequencies in a given window. Spectral flux measures the spectral correlation between successive windows and thus represents the degree of spectrum change between windows. Compactness sums over frequency bins of a Fast Fourier Transform (FFT), which indicates the noisiness of the signal. Root mean square (RMS) refers to the square root of the mean of squares of magnitude. It is used to measure the amplitude in a window. Fraction of low amplitude frames refers to the fraction of previous windows in which RMS is lower than the mean RMS. This feature indicates the variability of the overall amplitude. Zero crossing rate is the number of times the time domain signal passes the zero-level in a given window. MFCCs describe the spectrum in a given window. They are achieved based on a linear cosine transform of a log power spectrum on a nonlinear mel scale of frequency. LPC describes a spectral window in a compressed form. Because of its accurate estimation of relevant parameters, it is widely used for speech analysis, compression and synthesis. Beat histogram is often used to represent the feature of tempo. Method of moments refers to the statistical measures such as mean, variance, skewness and kurtosis.

Based on these audio features, songs are grouped into different clusters using the Simple K-Means clustering algorithm (Kanungo et al., 2000).

3.2.2 Preliminary Evaluation

We presented the interface of *MusicSim* to 36 regular users of music players. 21 of them ($\approx$ 58%) thought that the similarity-based view was very useful and 11 ($\approx$31%) considered it as acceptable. We further discussed *MusicSim* in detail with another three users. After a brief introduction of *MusicSim*, we let them freely play with it using both their own and others' collections (size ranging between 271 and 789, with an average size of 459 songs). The results of the discussions were quite encouraging. Concerning the visual representation, all participants considered cover art a useful visual assistance. Displaying genres by color coding was commented as helpful for building an overview of their own collections, as well as shaping a quick impression of others' collections. The participants generally appreciated the idea of similarity-based browsing, and agreed that such an interface could become more useful in practice, if combined with the conventional list-based views. Similarity-based placement was sometimes confusing, especially when the results of audio analysis were inconsistent with the users' expectations. In this case, the participants seemed rely more on the reliable criteria, such as artist and album. Our user test was conducted with only three subjects, nevertheless, it showed a generally positive attitude towards similarity as a complement to standard list.

3.3 Summary and Discussion

This chapter has presented our initial exploration based on similarity. We built our first two prototypes *PhotoSim* and *MusicSim* to support similarity-based browsing. Although our preliminary evaluation was conducted with relatively small numbers of users, we derived several substantial implications for interface design and the improvement of similarity-based browsing.

Both prototypes provide an **overview** of the entire collection, through which the user can jump directly to a specific similarity-based cluster. *MusicSim* includes cover arts and genre color coding as the **visual assistance** for the non-visual music perception. Keyword-based search helps to locate certain items quickly in a large collection. For photos which initially lack meaningful names, we expected that a convenient tagging functionality can improve the performance of explicit query. Both systems integrate the map-based view with the traditional interface concepts: *PhotoSim* employs the map representation below the level of time. In *MusicSim*, songs included in a playlist are displayed in both list and map view. The results of the preliminarily tests revealed a desire for a more close coordination between the novel and traditional interface elements.

To enable an **active control** over the layout, the user can manually reposition items and clusters, or trigger a re-clustering by adjusting the number of clusters or the similarity threshold. An "undo" functionality allows a **recovery from undesired operations**. In order to collect more efficient user feedback and to encourage the user's involvement, more elaborate interactions should be allowed beyond current simple user manipulations.

3.3.1 Improving Similarity-based Browsing

Due to the computational cost, content analysis in *PhotoSim* and MusicSum runs off-line and it is one limitation for the practical usage of our prototypes. To offer more flexibility, we would need simple but efficient content analysis running in real time.

The concept of similarity-based browsing is appealing. However, due to its low accuracy and the semantic gap between low-level features and high-level perceptions, the performance of content analysis is not satisfactory. The organization based on the overall similarity is sometimes unintuitive and hard to understand, especially with acoustic (hence non-visual) similarity. In this case users tend to rely more on the commonly used criteria, such as time for photos, artist and album for music. Due to its high computational cost, a performance which does not meet the user's expectations, and the user's preference for reliable criteria, we suggest offering **similarity on demand**.

Besides improving the algorithms for content analysis, there are two possibilities to enhance the similarity-based browsing. One solution is to **couple similarity tightly with the common used and reliable criteria**. **User interaction** can be the other compensation for similarity. As Kirk et al. (2006) suggested, over-automatic tools may confuse users and one solution is to offer middle-ground suggestions and let users choose based on their personal judgment.

Researchers in the field of Content-Based Image Retrieval (CBIR) have already realized the necessity to include the user in the loop by offering relevance feedback (Zhou and Huang, 2003): The user provides judgments of the retrieval results, and the system proposes more relevant results accordingly. The simplest way to providing relevance feedback is to ask the user to rate the retrieval results. A more precise approach is to let the user to select one of the system-segmented regions in the image, or to score the relevance of each region, such as in Blobworld (Carson et al., 1999). Elaborate user interactions enable a more precise definition of areas of interest. Fist (Cui and Zhang, 2007) allows two levels of user feedback: On the global level, the user can give scores to the retrieved image by clicking the "correct" or "cross" marks.

On the object level, the user can freely define relevant and undesired areas by drawing 2 different types of strokes respectively on any part of an image. A similar concept of applying similarity on local levels has been applied in Shades of Music (Baur et al., 2009), an example-based search interface which enables the user to discover songs containing similar parts of a focus song.

Concerning the browsing interface for media collections, we expected that applying **similarity on local levels** will be more understandable than on an overall level. For example, the user can define an area of interest in a photo or choose a representative audio clip, based on which the system can return more relevant items. To improve the user's active control, the relevant parameters, such as the similarity threshold, the number of clusters or the length of an audio sample, should be adjustable. Beside offering relevance feedback, we expected that additional information which cannot be extracted from content analysis, can be also derived with the users' explicit input, such as semantic relations reflected by the user-generated tags.

The next chapter presents our progress of combining similarity with other categories of information to enhance the *Exploratory Browsing* experience with personal media collections.

4

Combining Similarity with User-Generated Data

Our initial exploration in this thesis started with similarity and we developed two similarity-based browsing interfaces *PhotoSim* and *MusicSim* (cf. Chapter 3). The preliminary evaluation confirmed the usefulness of similarity in organization and browsing. However, similarity cannot efficiently reflect the high-level perceptions and users rely more on reliable criteria. We believe that similarity should be used as a secondary criterion and offered on demand. We also suggested improving the similarity-based browsing by tightly coupling similarity with reliable criteria and by integrating more elaborate user interactions. Besides a closer coordination between similarity and commonly used criteria, such as time for photos, and artist and album for music, we expected that the browsing experience can be further improved by introducing additional information, especially information that can be collected through user interactions.

This chapter presents our progress of supporting *Exploratory Browsing* with personal media collections based on a combination of three types of information: reliable metadata, similarity and additional information, specifically User-Generated Data (UGD). UGD in this chapter refers to the data generated by single users, such as item popularity extracted from the user's consumption records and tags explicitly given by the user. We developed two prototypes *CloudMonster* (Chen et al., 2009a) and *PhotoMagnets* to enhance the browsing flexibility with personal photo and music collections respectively. Beyond the simple manual override in *PhotoSim* and *MusicSim*, we proposed elaborate interactions to give the user more active control over the solutions generated by the system. The observation of the users' behavior in the evaluation helped to deepen our understanding of the *Exploratory Browsing* activities with personal media collections, especially the prominent *General Purpose Browsing*.

The findings gathered during the development and study of the four prototypes presented in this part helped to refine the initial model of *Exploratory Browsing* by uncovering the underlying transitional relations between *General Purpose Browsing* and the other two browsing activities.

4.1 CloudMonster: Enhancing Browsing Flexibility with Music Collections

Our preliminary test with *MusicSim* revealed that, users have a fundamental requirement for commonly used and reliable criteria, such as artist and album. The user study also confirmed the effectiveness of similarity to enhance the browsing flexibility. Besides similarity, some commercial products have already introduced additional information, such as User-Generated Data (UGD). For example, iTunes released a sidebar named Genius[9] which recommends songs that are compatible with the selected song based on both acoustic similarity of the music and the users' consumption records.

We believe that such a combination of similarity and additional information can improve the *Exploratory Browsing* experience with music collections, but that a more visual interface than Genius is needed. To collect the users' practical requirements of non-standard criteria, we first conducted a survey. Based on the results of this survey, we developed *CloudMonster* to enhance the browsing flexibility with music collections.

4.1.1 Survey

The objective of this survey was to investigate how users organize their music collections and their ways of browsing and searching. The key issues addressed were the commonly used and additional desired criteria. The questionnaire consisted of three parts: the demographic information, the usage of music players, and the behavior of music organization and playlist generation. We recruited 36 college students, 25 male and 11 female. Their ages ranged from 20 to 28, with an average age of 21.6 years.

According to the participants' multiple answers, the most popular music players were Winamp[10] (36.10%), Windows Media Player[11] (27.80%) and iTunes (22.20%). The main reason for choosing music player were user-friendless (41.49%), flexibility (21.28%), compatibility (17.02%) and price (13.83%). We also found some correlations between music player and the relevant user behavior. For example, iTunes offers multiple views of the entire collection and allows generating playlists based on the listening history. The participants, who were regular iTunes users, tended to use cover art and listening history more often.

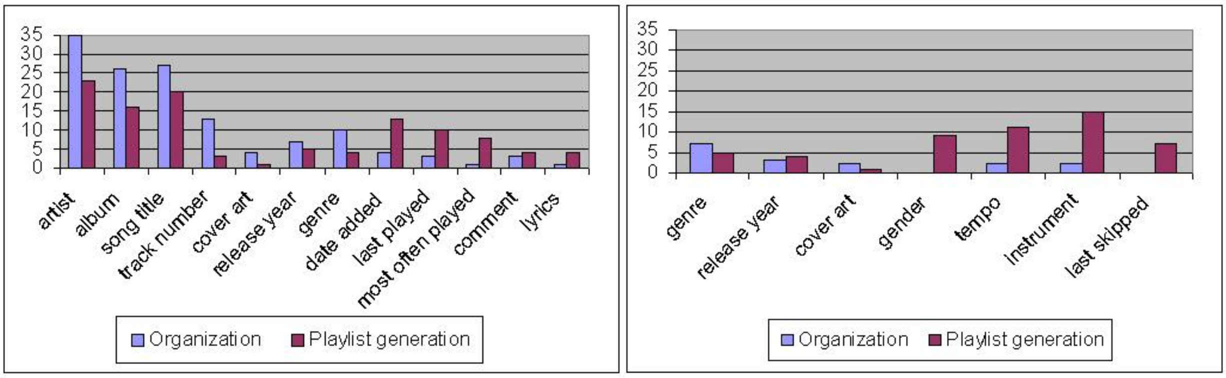
---

9 http://www.apple.com/retail/geniusbar

10 http://www.winamp.com

11 http://www.microsoft.com/windows/windowsmedia/default.mspx

Artist, album and song were the most common criteria for both organization and playlist generation (see Figure 4.1a). Other important criteria for organization were track number, cover, release year and genre. For playlist generation, listening history, comment and lyrics were also perceived as useful. Regarding the desired criteria, the most frequently requested attributes were gender, tempo and instrument (see Figure 4.1b).



(a) (b)

Figure 4.1: Results of the survey: a) Commonly used criteria. b) Desired criteria.

Based on the result of the survey, we decided to utilize the commonly used metadata such as artist, album and song, as the main criteria for organization and browsing, but also to include additional criteria, such as genre, gender and instrument, to enhance the browsing flexibility. We developed *CloudMonster* to support freely browsing of music collections.

4.1.2 Interface

CloudMonster is implemented using the Prefuse toolkit (Heer et al., 2005). As Figure 4.2 shows, the interface of *CloudMonster* consists of three columns: The main view and list view in the middle, the lists for initial views, color coding schemes and magnets on the right, and the panels for overview, search, saved playlist and search criteria, and genre histogram on the left. Songs are displayed both in the list and the main views. In the list view, songs can be sorted by id, genre or the title of artist, album or song. In the main view, songs are represented as either colored nodes or corresponding cover arts. With mouse hovering on one node, the details of this song, such as cover art, artist name and song title, is shown in a tooltip. One song can be played by double clicking, and the user can create a playlist for the selected songs by the “new playlist” button in the playlist panel.

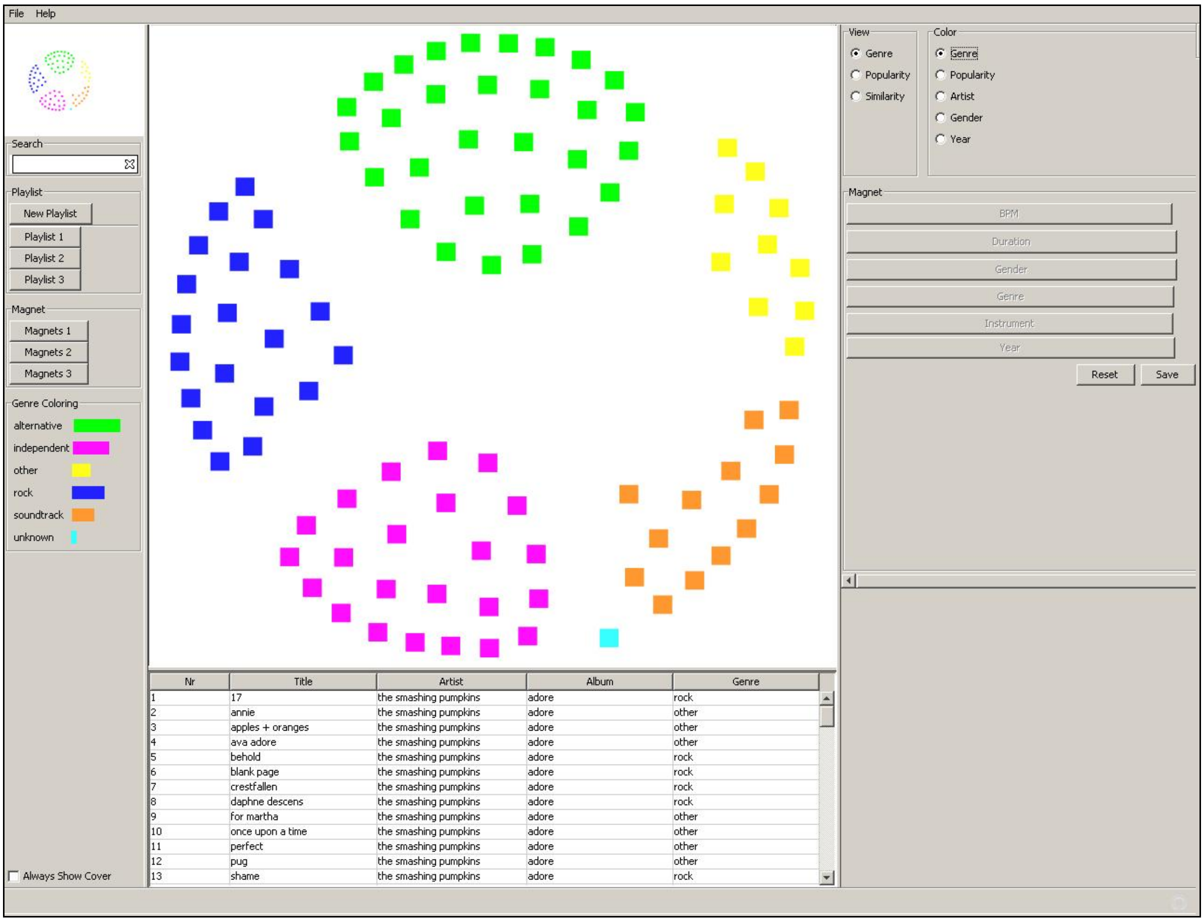
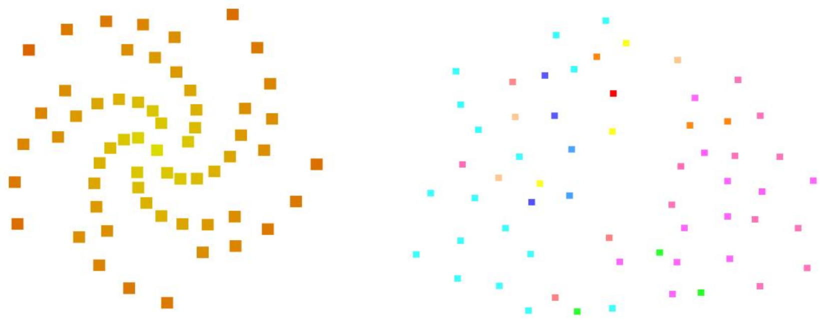


Figure 4.2: The interface of CloudMonster consists of three columns: A main view and a list view in the middle; the lists for initial views, color coding schemes and magnets on the right; the panels for overview, search, saved playlist and search criteria, and genre histogram on the left.

Initial Views

As users may desire additional criteria for organization and browsing, we offered 3 views for the entire collection. Genre conveys general information about music (Cunningham et al., 2006), and thus a genre view is provided, in which songs are grouped into genre-related sectors(see Figure 4.2). The user’s listening history, especially the popularity of songs, gives an overall impression of the popular set in the personal collection, and thus is provided in a popularity view (see Figure 4.3a). The positions of songs are determined by their popularity, which is retrieved from Last.fm[12], and the most popular songs are placed in the center. A similarity view is created based on the same method applied in *MusicSim* (cf. Underlying Audio Analysis in Section 3.2.1 in Chapter 3). In this view, acoustically similar songs are placed near to each other (see Figure 4.3b). The user can manually rearrange these views by drag-and-drop.

[12] http://www.last.fm



(a) (b)

Figure 4.3: Two initial views: a) The popularity view with color coding of popularity. b) The similarity view with color coding of artists.

Visual Assistance

Cover art is usually employed as a visual assistance for music collections, which is also included in *CloudMonster*. Moreover, 5 different schemes for node color coding are provided: artist, gender, genre, popularity and release year. For ordinal parameters, each color presents one category, such as gender, genre or artist (see Figure 4.2 and 4.3b respectively). For quantitative parameters such as popularity and release year, the transparency presents the popularity (see Figure 4.3a) or recentness. An additional genre histogram is provided to illustrate the genre distribution in the entire collection (see the bottom left panel in Figure 4.2). By double clicking on one genre name, the user can assign new color to each genre in an extra color panel.

Magnet-based Search

As users often change their mind during the search process (Bentley et al., 2006), a high degree of flexibility for query formulation and refinement should be supported. Non-expert users may have difficulties to express their musical preferences in a formal way (Vignoli, 2004), and an efficient assistance for query formulation beyond the traditional keyword-based search will be appealing. Inspired by Dust&Magnet (Yi et al., 2005) and Dynamic Query (Ahlberg and Shneiderman, 1994a), we apply a magnet metaphor to support flexible search. It allows the user to select a particular set of songs based on one or more criteria. Each criterion acts as a magnet and attracts matching songs while repelling others. We expected that this magnet metaphor can facilitate complex query formulation and offer an intuitive visualization for the retrieval results.

Based on the results of the survey presented in Section 4.1.1, 6 types of magnets are offered (see the bottom right panel in Figure 4.2): genre, BPM (beats per minute), duration, release year, gender and instrument. The corresponding information for the first four types is extracted from songs' ID3 tags. Gender and instrument are assigned manually, due to the lack of an appropriate analysis algorithm. By selecting a criterion in the magnet list, a magnet (a black square with the name of the selected criterion shown in the center) is automatically generated in the main view. The size represents its magnitude, which can be adjusted by mouse wheeling. The user can set the value for each magnet in the magnet list and the matching songs will be automatically attracted. Songs attracted by multiple magnets will be places between them as a circle. The user can freely adjust the position of each magnet and the relevant songs will be replaced accordingly. Figure 4.4 shows the results of a search based on three magnets, and one can observe that songs in the center circle satisfy these three criteria.

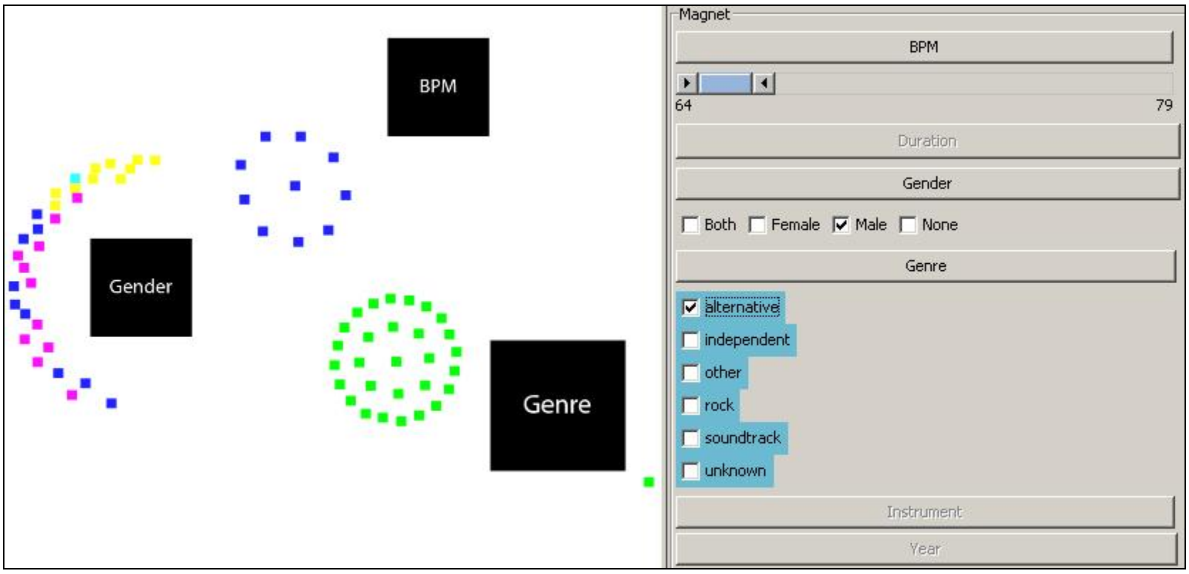


Figure 4.4: An example of magnet-based search.

Besides the 6 types of magnets, a similarity magnet is introduced. The user can select a song in the list view, and it will be highlighted in the similarity view, and automatically attract similar songs and repel others. Similar to a playlist, the visualization of the retrieval results can also be saved with the "save" button in the bottom left corner of the magnet list. The user can switch to a previous visualization by selecting a saved playlist or search criteria (see the middle two panels in Figure 4.2 on the left). Same to *PhotoSim* and *MusicSim*, the user can also use an "undo" functionality to withdraw the latest operation.

4.1.3 Evaluation

We conducted a user study to evaluate the performance of *CloudMonster*. One of our main concerns was the effectiveness of magnets in supporting different search tasks.

Settings and Procedure

The study was conducted in our laboratory and the participants were equipped with a laptop and a mouse. The participant was tested with a subset of their own collections with 120 songs. The study lasted on average about 45 minutes per participant. It was recorded on video using the Think-Aloud protocol. All the scores in the questionnaires were rated on a 11-point Linkert-scale where 10 represented the highest score.

At the beginning of the study, the participants were asked to fill out a pre-questionnaire with demographic information and general experience with music collections. In order to confirm the important search criteria, they were asked to execute two tasks of mentally music recalling. They mentally recalled one favorite song and wrote down the criteria they would use to search for it, besides the names of artist, album and song. After this independent recall task, the participants were shown the magnet list in *CloudMonster* and asked to conduct the same recalling task again (assisted recall). After completing these two recalling tasks, they were given a brief introduction of *CloudMonster*. They played around with it in a 10-minutes trial, and then executed 4 search tasks: Locate the favorite songs, search for these favorite songs with higher tempo, search for songs similar to the song in the recalling task, and search for songs from the favorite genre and released in the same year. Then the participants were asked to generate a playlist for a party. They were first asked to describe the ideal party music with texts and then generate such a playlist in *CloudMonster*. After completing all tasks, they filled out a post-questionnaire concerning the overall impression of *CloudMonster*.

Participants

We recruited 12 participants, 7 female and 5 male. Their age ranged from 20 to 29, with an average age of 24.2 years. All participants were regular users of music players who own large music collections.

Results

User performance: Genre, year, instrument, gender and BPM were the frequently mentioned criteria in both independent and assisted recalling tasks.

The participants' behavior was overall consistent in all tasks. To locate the favorite songs in task 1, all of them conducted a keyword-based search. They commented that, scrolling in a long list or panning-and-zooming in a large graph is problematic, and the standard keyword-based search has an advantage in this case. To search for the favorite songs with higher tempo in task 2, all participants first located one favorite artist with keyword-based search, and then used a BPM-magnet to attract songs with higher BPM value from this set. One user even clustered all songs into three sub-groups by adjusting the value and position of the BMP-magnet for three times, which revealed a potential usage of magnets in supporting clustering tasks. To look for songs similar to one example song in task 3, the participants first switched to the similarity view, then chose one song in the list view. Songs similar to this song were automatically attracted to it in the main view. To search songs from the favorite genre and released in the same year in task 4, the participants used the genre- and year-magnets. 11 out of 12 participants, who placed the two magnets linearly, spontaneously understood that the matching songs were those in the circle between the two magnets. The other participant stacked the two magnets in one pile and then the matching songs were those attracted nearest to the magnets.

In the task of playlist generation, the most frequently mentioned descriptors were tempo, genre, popularity, mood and instrument. Two representative examples were: up-tempo, happy but no too happy, lots of guitar; songs everybody knows, not too long, different artists. Except mood, the other attributes were offered in *CloudMonster*.

Overall impression: *CloudMonster* received positive feedback in the aspects of helpfulness (M=8.75, SD=1.42), enjoyment (M=8.50, SD=1.31) and feeling of control (M=8.75, SD=1.22). Understandability was rated lower (M=7.50, SD=1.62), due to the inconsistency of the criteria used in the initial views, color coding schemes and magnet list. For example, some participants were confused that popularity appeared in the initial view and color coding but not in the magnet list. The magnet metaphor was appreciated (M=8.58, SD=1.00). The initial views and the color coding schemes were commented as useful for deriving an overall impression of the entire collection. The initial views were highly appreciated: genre (M=8.75, SD=1.54), popularity (M=8.25, SD=1.22) and similarity (M=7.75, SD=1.86). The similarity view was scored lower, because the results of the content analysis were not satisfactory. Sub-structures were desired in the initial views. For example, in the genre view, the participants wished songs in each genre sector to be further grouped by artist or album. The color coding schemes were also appreciated: genre (M=8.25, SD=1.06), popularity (M=8.17, SD=1.40), gender (M=7.92, SD=1.62), artist (M=7.83, SD=1.85) and release year (M=7.13, SD=1.76). The participants desired more information about their playing histories rather than the release year. Besides the genre histogram, legends for other color coding were also desired.

Feedback on magnets: The magnet metaphor was perceived as a very helpful assistance for complex query formulation, and we also received valuable suggestions for improvement. Besides the offered 6 types of magnet, some extra criteria were desired, such as mood and language. Currently only one example song is allowed, and clustering task can be achieved by adjusting the value and location of a magnet for multiple times. We believed that allowing duplicate magnets and multiple example songs will bring more flexibility.

Another main concern was the instability of the visualization. We employed a force-directed layout approach to simulate magnetism. The distance between each pair of items is controlled by an attracting force, which is determined by their similarity. For items which are attracted very close to each other, a repulsive force keeps a certain distance between them, in order to achieve a non-overlapping layout. This method is overall easy to use and has already been applied in existing toolkits, such as Prefuse (Heer et al., 2005). However, it has one main shortcoming: Due to the co-existence of the attracting and repulsive forces, the stability of the layout decreases enormously with increasing numbers of similar items. Especially when large numbers of matching items are displayed in a circle, the visualization becomes very oscillate. Therefore, a more stable and visually appealing model for magnetism simulation should be introduced.

4.1.4 Discussion

We developed *CloudMonster* to enhance the browsing and searching flexibility with music collections. It received overall positive feedback in the evaluation, especially with the magnet metaphor. Its understandable metaphor enables easy query formulation and the active steering was commented as intuitive and easy to handle.

We offered **similarity on demand** and the user can browse in the similarity view or conduct an example-based search. Moreover, we combined similarity with reliable criteria (e.g., artist, album, duration and release year) and additional criteria (genre, BPM, gender, instrument, etc.). Song popularity extracted from the user's listening records was also included to offer an overview of the popular set in the collection. The participants' positive feedback confirmed the **importance of listening history** in music collections. Due to the lack of relevant feature extraction algorithms, the information about gender and instrument were assigned manually. We expect that a convenient tagging functionality can enhance the system performance by taking usage of user contributions.

Multiple representations and relevant **color coding schemes** of the entire collection can help the user forming an overall impression of the collection. In each representation, songs can be placed in different ways, such as displayed on a map or in a spiral, or grouped into several sectors. For sectors such as in genre view, songs should be further grouped into subsets for example based on artist or album.

Magnets are efficient assistance for complex query formulation. In the magnet representation, with active control, such as dynamically adjusting the position and magnitude of a magnet, the user can understand the underlying system logic and create personalized visualization. The original magnet metaphor is enhanced to alleviate the problems of occlusion by automatically placing songs matching multiple magnets in a circle. To enhance the reproducibility, the user can save the arrangement of the visualization. An "undo" functionality enables the user **recovering from an unexpected action**. The current magnet metaphor needs further improvement. Duplicate magnets and multiple examples should be allowed. A new magnetism simulation approach should be introduced to solve the instability problem of the current method of force-directed layout.

In *PhotoSim* and *MusicSim*, changes caused by the user operations cannot be reflected in real time and the user's interaction and the system's response is more alike a turn-taking facility. Magnets in *CloudMonster* improve the **continuity of the user experience** by enabling a synchronous feedback with the user interaction, which improve the user's seamless experience, as suggested by Dynamic Query (Ahlberg and Shneiderman, 1994a). The visualization for the retrieval results was automatically generated and cannot be overridden. More active user control will help creating more satisfactory and personalized representations.

Regarding the user study, we conducted an **empirical evaluation** with *CloudMonster*, in which the participants were asked to conduct several pre-defined tasks. It helped us to evaluate the main functionalities, especially the magnet-relevant features. However, the users' activities were constrained by the specific tasks, and consequently we could not collect detailed insights about the users' practical behavior. Therefore, we expected to gain more substantial information about the practical usage of such a prototype through an explorative study. Including open tasks will enable a closer observation of the users' practical activities.

Supporting Exploratory Browsing

A standard list and keyword-based search are especially useful when search for specific items in a large collection. Magnets are intuitive and efficient assistance for complex query refinement. Illustrating information, such as artist and album, on different levels of the hierarchical structure is essential to keep the most reliable criteria always in track. *General Purpose Browsing* is less effectively supported in *CloudMonster*, as there is no option for the user to explicitly define an area of interest within the whole collection. Although magnets can be used to find certain items in the entire collection, they may destroy the original representation. Thus the initial views should be separated from the panels where unstructured activities, such as magnet-based search, may take place.

4.2 PhotoMagnets: Enhancing Browsing Flexibility with Photo Collections

As confirmed in *CloudMonster*, combining similarity with additional information can effectively enhance the browsing flexibility. Integrating User-Generated Data (UGD), such as song popularity extracted from the user's listening records, can bring more browsing options. The magnet metaphor enables high flexibility for complex query formulation, and the corresponding interactions enable an active user control and seamless experience, and thus motivate the users' involvement. Encouraged by the positive feedback received towards *CloudMonster*, we explored how to enhance browsing flexibility with photo collections. We developed *PhotoMagnets*, a prototype employs the magnet metaphor in addition to more traditional interface elements to support a flexible combination of structured and unstructured photo browsing.

4.2.1 Initial Concept Development

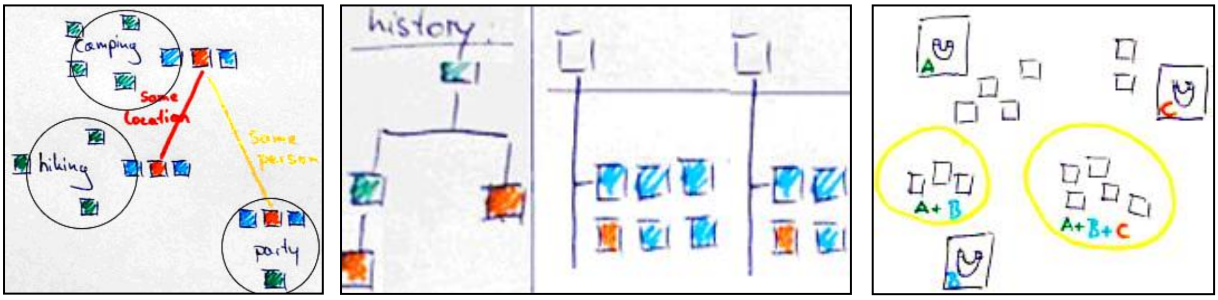
People's activities around their photo collections are often highly dynamic and unstructured, such as casual browsing and searching, or loosely structured storytelling (Balabanović et al., 2000; van House, 2004). Today's interfaces with fixed and pre-defined structures cannot efficiently support the flexibility and dynamics observed in human behavior. In general we believe that an interface supporting flexible activities should follow these common characteristics:

Encourage exploration: As discussed in Chapter 2, people tend to spend more time in browsing than in explicit search. Therefore, it is essential to support exploratory activities beyond direct search. As users may switch frequently between different browsing activities, they should be integrated seamlessly and tightly in the interface.

Introduce serendipity: On the way of exploration, users may easily get sidetracked by serendipitous findings. The system should facilitate serendipity and make it easy to revisit old collections and to rediscover forgotten items. An anchor point should always be kept on the original exploration track, to enable a recovery from distractions.

Enable active interactions: As van House et al. (2004) discovered, users prefer interaction and feedback over full automation. Active interactions can help users gaining a better understanding of the underlying system logic and enhance their feeling of control and enjoyment.

Based on these general considerations, we developed two interface concepts for browsing and one for search, in which the user can actively manipulate the layout through interaction. Since people tend to keep photos organized around the same event (Rodden and Wood, 2003), we employed event as the basic criterion for browsing and searching. BubbleBrowser is a circular layout (see Figure 4.5a), in which each event is displayed as a bubble of representative photos. Other events sharing the same contexts, such as location or person, can be opened in other bubbles. A connecting line between two bubbles illustrates the logical relation between events and helps tracking the user's browsing history. Similar to BubbleBrowser, the browsing history in TreeBrowser is shown as a tree in an extra panel, and corresponding events are displayed by their browsing order (see Figure 4.5b). As the magnet metaphor was proved in *CloudMonster* efficient search assistance, we integrated it in *PhotoMagnets* (see Figure 4.5c): Multiple search criteria act as magnets, automatically attracting matching photos and repelling others. Other photos sharing common tags are suggested as semantic "related photos".



(a) (b) (c)

Figure 4.5: The sketches of the initial concepts (Reiter, 2009): a) BubbleBrowser. b) TreeBrowser. c) Magnet-based Search.

We made sketches for these concepts and tested them with 8 regular photo consumers. After introducing each concept with the sketches shown in Figure 4.5, we illustrated them with an example collection of 90 print photos. The results showed that both bubble and tree layouts were generally appreciated. The magnet metaphor and “related photos” were also perceived as useful. Although the connecting lines in BubbleBrowser illustrated the browsing history, the multiple connected bubbles seemed confusing especially when the connected events were topically irrelevant. TreeBrowser was perceived as similar to the common way of hierarchical browsing, and 5 out of 8 participants preferred a chronological structure of the collection over the browsing history.

Based on the participants’ feedback, we refined our concepts: The connecting lines between bubbles were excluded from the BubbleBrowser. Since the structured browsing style in TreeBrowser was preferred with a clear chronological order, the hierarchical structure of the collection was shown on a separate timeline, and corresponding events were ordered by time. In order to improve the search flexibility, magnets could be dragged over the timeline to attract matching photos from specific events. To save the visualization of the retrieval results, a history sidebar was introduced in the magnet view. We conducted a second round of tests with 8 more participants and all of our refinements were confirmed. Overall, TreeBrowser was considered useful with a clear chronological structure, and BubbleBrowser and the magnet metaphor were appreciated for supporting flexible activities.

4.2.2 Interface

Our preliminary tests illustrated that chronological and loosely structured views have respective advantages in supporting structured and unstructured activities. Therefore, we decided to employ TreeBrowser for structured browsing, while BubbleBrowser and the magnet metaphor were merged into one view to support flexible browsing and searching. *PhotoMagnets* is implemented based on Piccolo toolkit (Bederson et al., 2004) for Zoomable User Interface design.

TreeBrowser

In TreeBrowser, the user can organize, browse and tag photos in a hierarchical event structure (see Figure 4.6). An event detection is automatically applied to the imported collection. By default, the first photo of each event is chosen as the event representative photo (RP), and the first event RP is chosen as the collection RP. All the imported collections are displayed chronologically on a scrollable timeline: Each collection is shown with its collection RP, accompanied by general information, such as a name and location provided by the user, and the automatically calculated time span. On the top of the collection RP, events in this collection are displayed as a temporal histogram (see Figure 4.6): Each peak represents one event, with the width indicating the event’s time span and the height illustrating the number of photos in this event. Additional information about an event can be shown with mouse hovering on one peak: The collection RP is temporally replaced by the corresponding event RP, and a tooltip shows the event name and the number of photos in it.

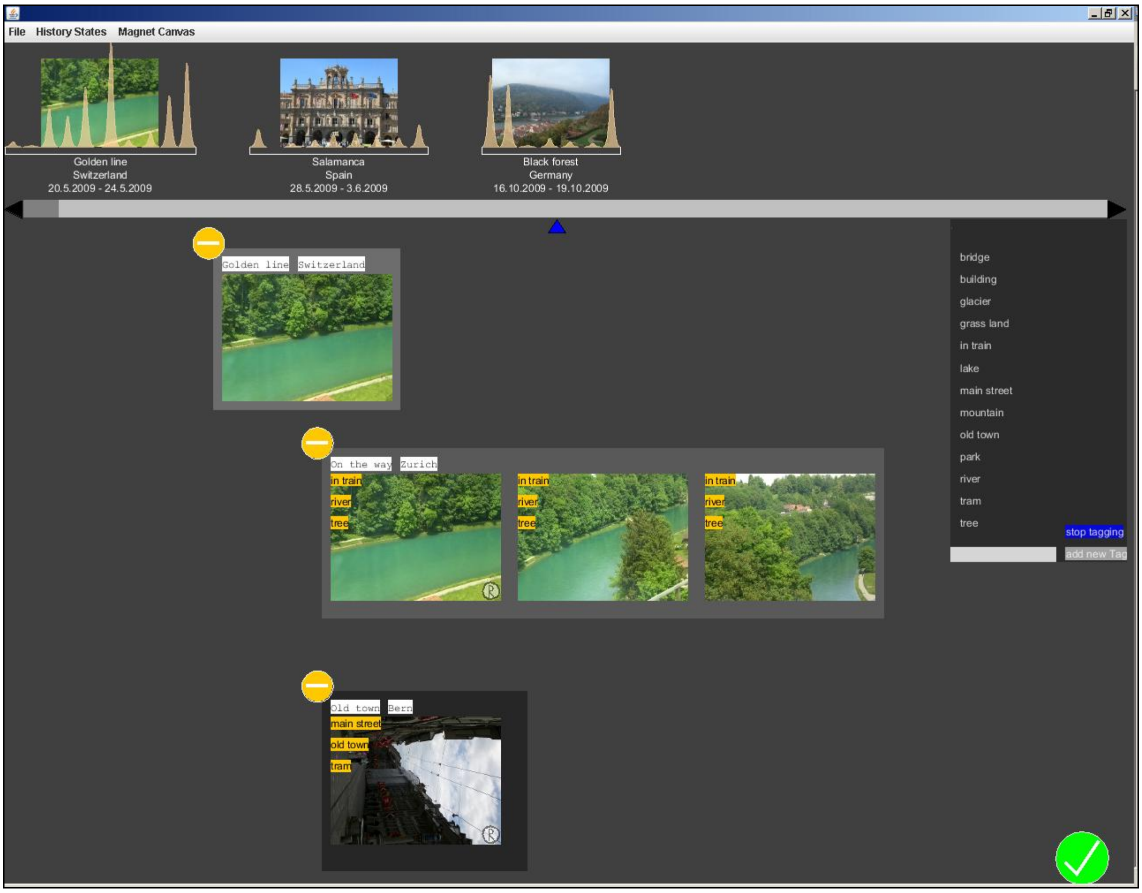


Figure 4.6: TreeBrowser consists of three components: A timeline (top), a presentation canvas (center) and a tag panel (right).

The user can load a collection onto the presentation canvas by dragging its collection RP from the timeline. As Figure 4.6 shows, the loaded collection is led by its collection RP, and photos are shown chronologically in respective event blocks. Each event RP is labelled with "R" in its right bottom corner. Through a shortcut menu triggered by the right mouse button, the user can set any photo as its collection or event RP. Above the collection RP and the first photo in each event, there are two text fields, in which the user can input the name and location for the entire collection or each event. The collection or event can be expanded or shrunk by clicking the yellow button on the top left corner of the collection RP or the event block. Clicking on one photo creates an extra slideshow window, in which all loaded photos are shown sequentially in full size.

TreeBrowser allows refining loaded events by drag-and-drop, and the chronological order is always automatically maintained. By dragging a photo into any area of a preceding event, this photo and all photos before it are automatically attached to the end of this event. Similarly, a photo and photos after it can be inserted at the beginning of a subsequent event. Two successive events can be merged by dragging the first photo of the first event into the second one, or by dragging the last photo of the second event into the first one. By dragging a photo to the blank area between the current event and its neighbour, a new event with corresponding photos is automatically inserted in between.

To improve the query efficiency with photos, which initially lack meaningful names, a tagging functionality is provided in TreeBrowser. New tags are generated by typing text into the bottom text field of the tag panel and then clicking the “add new tag” button. The user can select multiple tags and assign them to multiple photos by clicking on each of them once. Assigned tags can be removed by clicking them once. All changes can be saved by the green button in the bottom right corner of the canvas.

BubbleBrowser

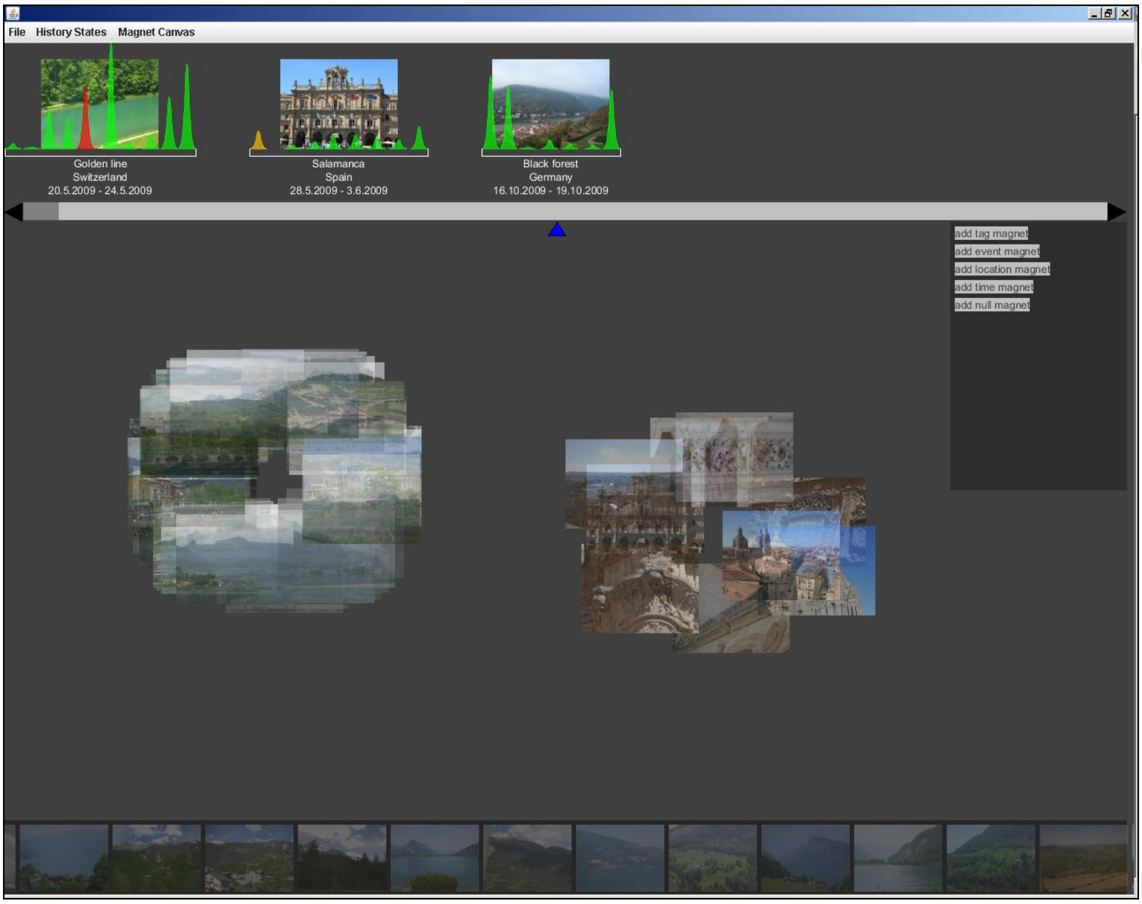
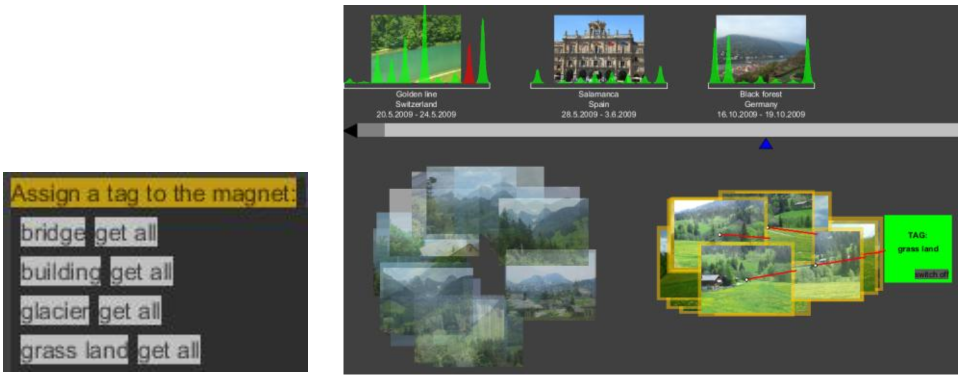


Figure 4.7: BubbleBrowser consists of four components: A timeline (top), a presentation canvas (center), a magnet list (right), and a panel for photo suggestions (bottom).

The user can switch between TreeBrowser and BubbleBrowser with the Tab key. BubbleBrowser is designed to support flexible browsing and searching. It consists of four panels (see Figure 4.7): a timeline at the top, a presentation canvas in the center, a magnet list on the right, and a panel for photo suggestions at the bottom. Similar to TreeBrowser, photos are loaded from the timeline to the canvas by drag-and-drop. They are displayed in event bubbles, ordered clockwise by time. Other photos, which share at least one common tag with the loaded photos, are displayed in a flow animation in the bottom panel for "photo suggestions". By dragging one suggested photo onto the canvas, a menu pops up and the user can choose between loading this photo only, or loading photos with same tags from the same event or collection. Consistent with TreeBrowser, double clicking on one loaded photo leads to the extra slideshow window. A photo can be deleted with Del key. Through the shortcut menu with mouse right clicking on a photo, the user can remove the relevant event from the canvas, or view it in TreeBrowser.

For magnet-based search, five types of magnets are offered in the magnet panel: tag-, event-, location-, time- and null-magnet. The first four magnets can be used to search photos matching certain attributes, for example, containing certain persons, from the same event or location, or captured in a certain time period. To create a neat representation, the null-magnet is used to group photos that do not match any other magnets. By clicking one of the tag-, event- or location-magnets, relevant attributes, such as tags, event names and locations, are shown in a pop-up window. Figure 4.8a is an example for a tag-magnet. In contrast, a time-magnet provides four sliders, with which the user can define a specific time period by setting a seed date (year, month and day) and a time span. There is no pop-up window for null-magnet.

Magnets can be used to attract matches from loaded photos, from all collections, or from certain events. In the first case, if the user has a clear idea about the target events, she or he can first load them on the canvas. By choosing one of the attributes, such as a tag "grass land" in Figure 4.8a, the cursor becomes a magnet, a green square with the magnet type "tag" and the attribute "grass land" shown in the center. After dropping the magnet onto the canvas, the matching photos are gradually attracted to it and thus forming a new cluster (see the highlighted cluster in Figure 4.8b on the right). The red line(s) illustrate by which magnet(s) each photo is attracted.



(a) (b)

Figure 4.8: Using a magnet to attract matches from loaded photos. a) Setting the magnet attribute. b) After dropping magnet on the canvas, it attracts matches from loaded photos.

Events in the timeline are rendered differently according to the number of photos loaded from them: Green for none, yellow for some and red means all photos from this event have been loaded (see also Figure 4.7, 4.9, 4.10). The magnet's magnitude is reflected by its size and can be adjusted by mouse wheeling while pressing the Alt key (see the two magnets in Figure 4.11 with different sizes). The position of a magnet can be changed freely by drag-and-drop and photos attracted by this magnet will be replaced accordingly. A magnet can be activate/deactivated by clicking the "switch on/off" button on it. A magnet can be deleted by choosing "delete" in its shortcut menu with mouse right clicking.

In the second usage case, the user can use magnet to attract all matches from the timeline by choosing "get all" in Figure 4.8a. After dropping the magnet onto the canvas, matches outside the loaded event are automatically loaded and form another new cluster (see the one on the right hand side in Figure 4.9).

Figure 4.9: Using a magnet to attract all matches from the timeline.

Besides attracting loaded photos or loading all matches, magnet can be used to attract photos from certain events in the timeline. In the shortcut menu opened by a right click on a magnet, the user can choose "swipe timeline" and the cursor becomes a magnet. Events and collections containing matches are highlighted (see Figure 4.10). The user can swipe the magnet over the timeline, which has the same effect as normal mouse hovering. The user can select events containing matches by clicking on them, which will be highlighted in red and the number of total matches will be shown in the magnet. By dropping the magnet back on the canvas, matches from the selected events are loaded, similar to Figure 4.9.

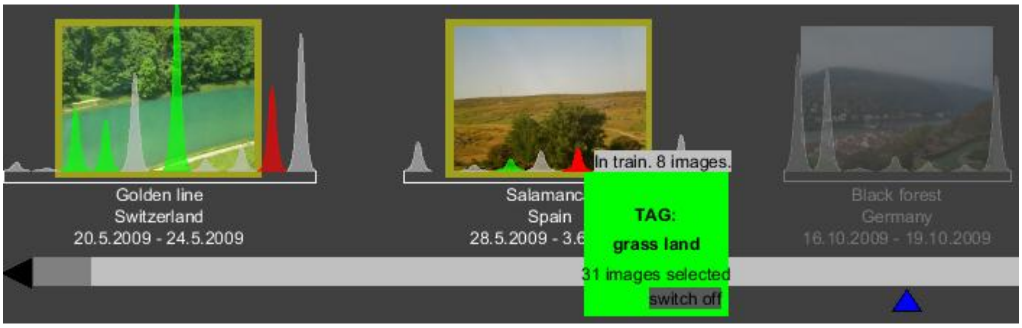


Figure 4.10: Using a magnet to attract matches from certain events in the timeline.

Besides the five types of magnets, there is another type of similarity-magnet. In the shortcut menu of a photo, the user can choose “convert to similarity-magnet”, and this photo will be used to attract visually similar photos. By defining multiple or duplicate magnets, the user can conduct multiple-magnets-based search. Figure 4.11 illustrates the retrieval results of a location-magnet, a time-magnet and a similarity-magnet (highlighted with a green frame). Photos matching multiple magnets will be placed between the magnets.

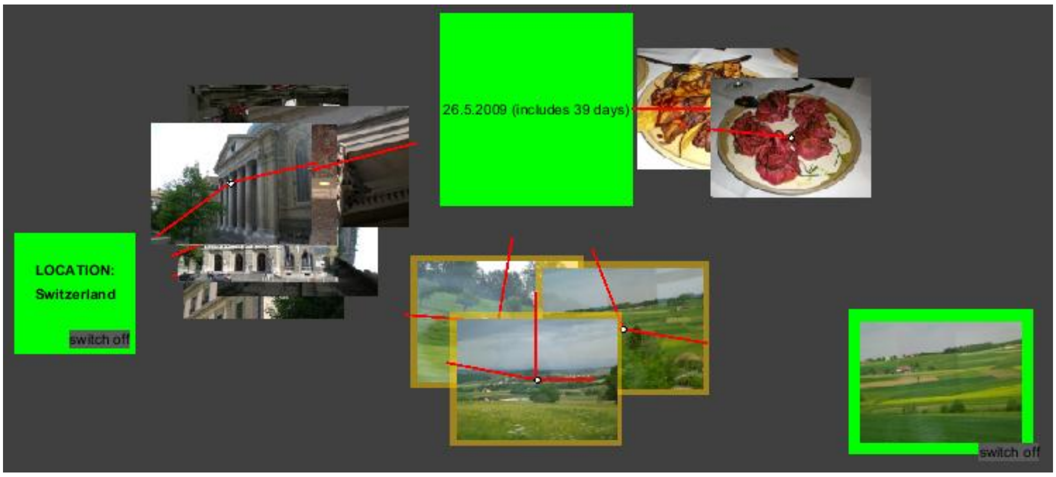


Figure 4.11: An example of magnet-based search: 3 tag-magnets and 1 similarity-magnet are used.

A snapshot of the current state in BubbleBrowser is automatically created when a new magnet is added in. The user can also create snapshots manually with the "history states" menu. All the snapshots are displayed in an extra history sidebar (shown/hidden with the H key), in which the yellow frame indicates the current screen (see Figure 4.12). The user can switch to a previous status by selecting one snapshot.

Figure 4.12: The history sidebar holds all saved snapshots, with the yellow frame indicating the current screen.

Underlying Algorithms

Event detection: Similar to Phototoc (Platt et al., 2002), we derived events as follows (see Fulmar 4.1): For all chronologically sorted photos, the event boundary is defined as a noticeable gap in the capture time. If g_i is the time difference between photo i and i+1, g_N is considered a gap if it is larger than a local average of gaps on a logarithmic scale. Empirically, the threshold K is chosen as log(17), and the window size d is set to 10. If N+i exceeds the size of the collection, the term is ignored.

$$\log(g_N) \geq K + \frac{1}{2d+1}\sum_{i=-d}^{d}\log(g_{N+i}) \tag{4.1}$$

Image analysis: Image similarity is calculated based on the LIRE framework (Lux and Chatzichristofis, 2008), a Java Library supporting several feature extractors. After comparing the accuracy and efficiency with a collection of 300 photos, we chose the FCTH feature extractor to calculate the similarity between photos. FCTH includes color and texture information in one quantized histogram (Chatzichristofis and Boutalis, 2008). A 10-bin histogram is first derived for 10 colors in the HSV color space, which are pre-selected based on the positions of the vertical edges in each channel. By separating each color in 3 hues of dark color, color and light color, the histogram is extended to 24 bins. Haar Wavelet transform is applied to the luminance component Y in YIQ color space. The feature histogram is then enriched by the extracted textural information and expanded to 192 bins. A 3bits/bin quantization is applied to limit the final length of the feature descriptor to 576 bits per image. The Tanimoto coefficient (Chi et al., 1996) is chosen to measure the feature distance between images, namely their similarity. It is then normalized to a value between 0 and 1.

Figure 4.13: A similarity-magnet (highlighted with green frame) with the attracted visually similar photos.

When a photo is converted to a similarity-magnet, the similarity between this example photo and other loaded photos will be dynamically calculated. Photos with similarity value exceeding a threshold (empirically chosen as 0.5) will be included in the retrieval results (see Figure 4.13).

Model for magnetism simulation: As discussed in CloudMonter, the popular force-directed layout approach is not suitable to display an interactive visualization with large numbers of items. Therefore, we propose another method to simulate magnetism. Instead of defining both attracting and repulsing forces for each pair of items, we only introduce the attracting force for each pair of magnet and matching photo. By pressing the left mouse button on one magnet, the attracting force is activated. Consequently, the matching photos are gradually attracted to their relevant magnets. Photos attracted by more than one magnet will be placed between these magnets (see Figure 4.11). The attracting force is equal for tag-, event-, location- and null-magnet. For those attracted by similarity- or time-magnet, the attracting force is directly proportional to their similarity with the example photo or their time difference with the time seed. The attracting force is canceled when the left mouse button is released or photos reach the boundary of the magnet.

For each matching photo, its overall movement $\vec{V}$ is calculated with Formula 4.2: n is the number of magnets attracting this photo. $\vec{L}_p$ is the photo location. For i-th magnet, $\vec{L}_i$ is its location, m_i is its magnitude (set to 1.5 by default) and a_i is its attraction force.

$$\vec{V} = \sum_{i=0}^{n-1} \frac{m_i * a_i}{|\vec{L}_i - \vec{L}_p|} * (\vec{L}_i - \vec{L}_p) \tag{4.2}$$

For a tag-, event- or location-magnet, a_i is set as 1. For a similarity-magnet, a_i is the similarity between the photo and the example photo (between 0 and 1). For a time-magnet, a_i is calculated with Formula 4.3. In this formula, △t is the time difference between the photo's capture time and the seed date of the magnet. t_{span} is the time span of the magnet.

$$a_i = 1 - \frac{\Delta t}{t_{span}} \tag{4.3}$$

To avoid instability and oscillation, we do not apply a repulsive force to each pair of items. Instead, a repulsing functionality is introduced to solve the occlusion problem on demand: By holding the space key, photos repulse each other, gradually forming a non-overlapping layout.

4.2.3 Evaluation

In order to assess the effectiveness of *PhotoMagnets* in supporting flexible browsing and searching, we conducted an in-depth evaluation, in which both a usability test and an explorative study were conducted. We were specifically interested in the performance of the two browsing concepts in supporting different browsing activities.

Settings and Procedure

Each participant provided three photo collections (with an average size of 150 photos). The time span of these collections ranged from 1 to 15 years. During the study the participants were equipped with a laptop. The study was conducted at the participants' home and the participants were equipped with a laptop. The study lasted on average about 2 hours per participant. It was recorded on video using the Think-Aloud protocol. All the scores in the questionnaires were rated on a 5-point Linkert-scale where 5 represented the highest score.

At the beginning of the study, the participants were asked to fill out a pre-questionnaire with demographic information and the general experience with their photo collections. After a brief introduction of *PhotoMagnets*, the participants played around with a trial collection of 149 photos. After this 30-minutes trial, the participants joined an interview. In the first session, the system usability was tested based on a set of pre-defined tasks, which covered the main functionalities of *PhotoMagnets*: In TreeBrowser, the participants imported their photos and refine the events suggested by the system if they felt it was necessary. Then they tagged their photos freely. Switching to BubbleBrowser, the participants executed a number of tasks covering the main functionality, such as loading photos from certain events, and conducting a multiple-magnets-based search. To encourage the participants to use *PhotoMagnets* more freely and to give us a chance to observe their behavior more closely, they were asked to execute two open tasks in the second session: Introducing their photos to the interviewer and creating a photo album as a gift for a friend who will move to another city. After completing all the tasks, they filled out a post-questionnaire concerning the overall impression of *PhotoMagnets*.

Participants

We recruited 18 participants, 6 female and 12 male. Their age ranges from 19 to 37 with an average age of 27 years. All the participants were regular PC users and hobby photographers who own large photo collections. Only 4 participants had assigned tags to their photos and none of them had tagged more than 30% of their collections.

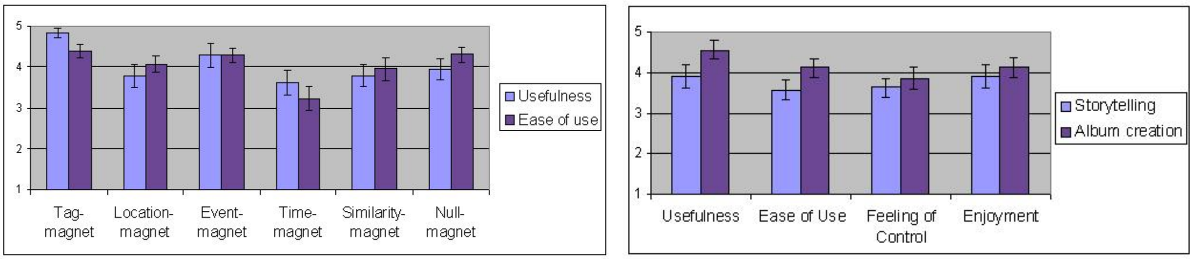
Results

During the evaluation we received a generally positive feedback to *PhotoMagnets*. The main functionalities, such as event segmentation, tagging, and loading of photos from the timeline, were generally appreciated. The magnets were especially appealing in facilitating search, clustering and comparison.

System usability: The timeline was rated as very useful (M=4.61, SD=0.50) and considered an important element that offers general information about the collections, such as time span and photo distribution. Loading photos from the timeline by drag-and-drop was appreciated (M=4.72, SD=0.57). The color-coding of the event histogram was also found appealing (M=4.06, SD=1.16).

In TreeBrowser, participants were overall satisfied with the performance of the automatic event segmentation (M=4.83, SD=0.38). The manual adjustment of event was commented as useful (M=4.78, SD=0.55) but seldom used. Although tagging is normally considered a tedious and time-consuming task (Bentley et al., 2006; Rodden and Wood, 2003), the tagging functionality in TreeBrowser was extensively used and was commented as very useful (M=4.94, SD=0.24), mainly because of its easiness to assign multiple tags to multiple photos. Although there were no quantity requirements, the majority of our participants assigned multiple tags to 80% of their photos. Three participants expected the tagging functionality to be included also in BubbleBrowser.

Swiping a magnet to load photos in BubbleBrowser was appreciated (M=4.22, SD=0.94). In multiple-magnets-based search, the red lines on photos were commented as helpful for illustrating the relations between magnets and photos (M=4.17, SD=1.15). Figure 4.14a shows the average ratings for each magnet type in the aspects of usefulness and ease of use. The tag-magnet received highest scores. The null-magnet was considered as useful in maintaining a neat display. The similarity-magnet was appreciated and also commented as entertaining. The participants wished to have a more functional time-magnet which would, for example, automatically align attracted photos chronologically.



(a) (b)

Figure 4.14: Ratings of magnet performance (Error bars represent ± standard error of the mean): a) Average ratings for each magnet type. b) Average scores for magnets performance in storytelling and album creation tasks.

Observation of user behavior: In the second part of the interview, the participants' behavior was observed with two open tasks of storytelling and album creation. The tag-magnet was the most frequently used magnet type (used 4 times per participant on average). Event-, similarity- and null-magnets were used less and nobody used location- and time-magnets. Although the history sidebar was considered as useful, it was used by one participant, who had accidentally deleted some photos. This illustrated that the history also represents an easily understandable "undo" functionality for interaction mistakes in addition to its original function of showing and conserving the exploration history as an anchor point back onto the original track. Considering the overall positive feedback and the limited size of the tested collections, we expect these functionalities to be used more often in a long-term usage. Photo suggestions were generally appreciated. During the two open tasks of storytelling and album creation, 2 and 5 participants respectively were sidetracked by the suggested photos. Regarding the performance in supporting these two tasks, magnets were appreciated in general and received higher scores with album creation (see Figure 4.14b).

We observed the participants' behavior during these tasks and derived the following results. The **storytelling behavior** was studied based on the photo- and story-driven modes proposed by Balabanović (2000) (cf. Section 2.1.5 in Chapter 2 for more details).

When asked to introduce their photos, 5 out of 18 participants started with TreeBrowser. The photo-driven mode dominated and they told the stories in the slideshow by explaining photos sequentially. With the clear chronological order, the story lines were fluid and the quality of the stories was generally high.

The other 13 participants started with BubbleBrowser. The story-driven mode was prominent: They told their stories based on selected photos, which were chosen among loaded candidates. Two common loading patterns were observed:

- Structured initial layout: Multiple bubbles were created either by loading certain events or only qualified photos with magnets. The bubbles were positioned with certain logical structures, for example the chronological order.
- Unstructured initial layout: All photos were loaded at once, without any structure.

After loading photos, two different strategies were applied to reposition photos in order to achieve a clearer story line:

- Magnet clustering: Magnets were used to group photos into different piles.
- Manual repositioning: Instead of using magnets, participants repositioned photos manually.

7 out of 13 participants started with an unstructured initial layout and used magnets to group photos into different topics. The other 6 started with a structured initial layout, 4 with manually repositioning and the other 2 with magnets. Figure 4.15 shows two stories created by the participants.

(a) (b)

Figure 4.15: Two stories created by the participants (Reiter, 2009): a) With manual repositioning. b) With automatic magnet clustering.

During storytelling, the participants demonstrated some creative usages. Besides searching and clustering, magnets also helped to derive implicit contextual information. For example, participant 9 used two tag magnets to attract photos containing two friends and the results reminded him of their relationship:

This was the first time they met. Obviously they were attracted by each other and they always appeared in a pair in these photos.

Participant 8 even discovered a new way to exclude undesired photos: By reducing the magnitude of a magnet below zero, a negative magnet was created, which repulsed instead of attracted photos. Having discovered this additional feature, this participant created more negative magnets.

In general, stories told based on a structured initial layout were the most interesting ones and the clear structure helped telling the story more fluidly. Storytelling in TreeBrowser especially benefited from the chronological order in maintaining the temporal context of the story, and thus was commented as easier than in BubbleBrowser.

All participants conducted the **album creation** task in the BubbleBrowser. Most of them first loaded some candidate photos with magnets, from which they made their final decisions, and then manually repositioned these selections as a collage. Different from storytelling, only structured initial layouts appeared with photo loading. Two strategies were applied to make selections from the loaded photos:

- Rejecting selection: Final selections were made by deleting unwanted photos (13 participants).
- Accepting selection: Desired photos were directly selected (5 participants).

Magnets were commented as useful for photo clustering and comparison. Although similarity-based search performed not always correctly, participants still considered it very interesting and some of them even got inspiration from the unexpected results, which is consistent with Toms' suggestion of inducing serendipity via poor similarity (Toms, 2000a). For example, participant 7 used a similarity-magnet to choose one photo from a set of embarrassing photos of one friend:

It implies that my friend looks alike this puppet. That is funny. I'll try some more.

Expert Interview

Besides testing *PhotoMagnets* with hobby photographers, we further discussed it with four experts, three employees in digital media press and one professional photographer. In general, they agreed that *PhotoMagnets* could also support their professional photo workflows. Specifically, they highlighted its value in supporting their routine work of searching with an unspecific goal and producing different draft layouts. The discussion with their clients could be enhanced with dynamically generated layouts. Magnets, especially the similarity-magnet, were highly appreciated for grouping photos with a feeling of visual progression. More advanced visual features were desired, such as searching for photos with a certain color theme and an adjustable similarity threshold. This means that some modifications will be necessary for introducing *PhotoMagnets* into the professional field, for example, introducing different organization principles beyond time and events.

4.2.4 Discussion

Similar to *CloudMonster*, we offered **similarity on demand** and the user can search photos visually similar to example photos. Similarity is calculated in real time based on an efficient image analysis algorithm. Besides the overall similarity, content analysis can be used to make content-based photo suggestions, or to provide more detailed content features, such as background color and light condition in photos. The browsing flexibility can be improved by combining similarity with time and **User-Generated Data (UGD)**. UGD used in *PhotoMagnets* are user-generated tags and the browsing history saved in the history sidebar.

PhotoMagnets offers **different representations** of the collection, and events can be displayed in a hierarchical structure or in bubbles. The timeline with temporal event histogram offers a good overview of the entire collection. The magnet metaphor is used in addition to traditional and well-known interface elements such as a tree view. Both structured and unstructured browsing styles are supported in the two respective views. A **semantic coordination** between them, such as using the loaded photos as the anchor points, can facilitate smooth switching and avoid losing track of the context. A convenient **tagging functionality** is offered in TreeBrowser, which was frequently used by the participants in the evaluation. User-generated tags help to enhance the performance of magnets and to illustrate underlying semantic relations between photos. In addition tagging was requested in BubbleBrowser.

As the results of the study confirmed, the **magnet metaphor** is consistent with the common sense of physical magnetism and illustrates the underlying system logic well. The interactive nature of the interface enables the user to execute complex tasks with simple and active control, and thus facilitates complex query formulation and refinement. Although the magnet metaphor is a rather novel concept, users seem to understand it and spontaneously discovered some advanced usages beyond the direct search, such as clustering and grouping, comprehending the implied social information, and the discovery of negative magnets. The participants were able to use magnets to accomplish the tasks of storytelling and album creation, and some of them spontaneously discovered several advanced usages of magnets, which confirmed the enabled a high degree of flexibility.

Comparing with *CloudMonster*, the magnet metaphor is further improved. A new type of null magnet is introduced to create a neat representation by grouping photos not attracted by any magnets. Duplicate magnets and multiple example photos are allowed. A magnet can be activated or deactivated, and the red lines on photos indicate the relation between photos and magnets. Besides loading matches from the whole collection, the user can use magnets to load matches from the loaded photos or from certain events. When swiping a magnet over the timeline, the events will be color rendered differently and the number of matches in a selected event will be shown in the magnet. This information gives the user a prediction of the retrieval results. The stability of the visualization was one of the main concerns, but it is successfully established using our specific method for the magnetism simulation. Different from automatically replacing matching items in a circle in *CloudMonster*, we solved occlusion in

PhotoMagnets by offering a repulsing functionality on demand. The **history sidebar** makes it easy to switch back to a previous status by reproducing an exact same representation. Besides manually saving the current status for future references, *PhotoMagnets* generates such a record automatically when a new magnet is added in, which reduces the user's effort spend for the history creation.

We received encouragement as well as valuable suggestions for magnet improvement. The most frequently requested feature was the tracking of time with magnets, for example, chronological alignment of photos attracted by a time-magnet. Tag-magnets are frequently used, so we expect that a well organized structure of tags will further facilitate tagging for example by offering efficient tag suggestions. Since the display will be influenced by each newly added magnet, a selection area is desired to exclude selected photos from the influence of magnets. The history sidebar enables reproducing a saved representation. However, temporarily regenerating an exact same layout for the same set of magnets cannot be promised due to the randomness of automatic repositioning. Thus the reproducibility of magnets should be improved. The size of the test set in the evaluation was rather small and the scalability issue needs further investigation towards realistic collection sizes.

Concerning the evaluation approach, **a combination of empirical and explorative studies** was applied. The system usability was tested with a set of pre-defined tasks, which covered the main functionalities of the system. To allow the participants to use the system freely and to give us a chance to observe their behavior more closely, we included two open-ended tasks of storytelling and album creation. The participants were better motivated to explore the tool extensively and discovered several new usages especially with magnets, which we did not expect in the beginning. Due to the limited size of the test collections and time constraint in the laboratory evaluation, we wish that a long term study can give implications for other less used functionalities, such as other magnet types besides tag-magnet, photo suggestions and the history sidebar. Besides evaluate *PhotoMagnets* with hobby photographers, we further discussed it with expert users, and their generally positive feedback indicated a potential usage of *PhotoMagnets* in a professional user group.

Reflection on the Model of Exploratory Browsing

The observation of user behavior in the two open tasks helped us to derive substantial insights on the model of *Exploratory Browsing*, especially the transitional relations between different browsing activities. In accordance with other studies of user behavior **(**Cove and Walsh, 1988; Kirk et al., 2006**)**, we noticed that users seldom conduct explicit search with well-defined targets, and instead often navigate through the collections. On the other hand, because of the **overall familiarity** with their own collections, users often start with an area of interest, and seldom let themselves completely be led by serendipity. In general, users tend to rely more on *General Purpose Browsing* more than search and *Serendipitous Browsing*. Within a relevant neighborhood, they either browse in detail, or use magnets to group and compare relevant photos. During such an exploration, users may gradually formulate their needs, and then jump directly to the target in the hierarchical structure or use magnets to conduct a strict search. On the way of browsing or searching, they may be distracted by unexpected findings and then enter the phase of *Serendipitous Browsing*. **Serendipity** encourages users to discover new findings, revisit older collections and rediscovery forgotten items. Serendipity In *PhotoMagnets* is induced by similarity-magnet and tag-based photo suggestions. Although content analysis is not always as precise as expected, in some cases it stimulates serendipity indeed by introducing less relevant photos. User-generated tags indicate the underlying semantic relations among photos, which the user may not be aware of, and thus create another kind of serendipity. Considering the effectiveness of tag-based photo suggestions in inducing serendipity, we expect that similarity can also be used to generate content-based suggestions. With magnets, the user can dynamically produce new representations, which facilitate comparison and discovery of unexpected findings. When users feel bored or notice that serendipity may distract them far away from the original attempt, they may abort the *Serendipitous Browsing* and manage a **recovery from distractions**. For example, in *PhotoMagnets*, some users jumped to a previous status in the history sidebar, deleted some loaded photos, or cleared the entire canvas and started a new session.

5

Discussion

In this part we explored supporting the *Exploratory Browsing* experience with personal media collection. Our initial exploration started with similarity and we developed two similarity-based browsing interfaces *PhotoSim* and *MusicSim*. The preliminary evaluation of these two prototypes confirmed a general positive feedback towards similarity-based browsing. However, content analysis does not perform as accurate as expected, and the high-level perceptions cannot be efficiently represented by the low-level features. Therefore, similarity is more suitable as a secondary criterion for organization and browsing. We suggested enhancing the performance of similarity by a tight coupling with more reliable metadata and by introducing more elaborate user interactions. Building upon the gathered insights on *PhotoSim* and *MusicSim*, we developed two prototypes *CloudMonster* and *PhotoMagnets*. We combined similarity with the other two categories of information: reliable metadata, such as time for photos, and artist and album for music; additional information of User-Generated Data (UGD), for example, song popularity extracted from the user's listening records in *CloudMonster*, and user-generated tags and browsing history in *PhotoMagnets*. To closely combine system intelligence with user feedback, and to encourage the user's involvement, we introduced more elaborate user interactions beyond the manually overrides offered in *PhotoSim* and *MusicSim*.

During the development and evaluation of these four prototypes, we derived substantial insights for interface design, evaluation methodology and the refinement of the model of *Exploratory Browsing*.

5.1 Considerations for Interface Design

To improve the information abundance, **three categories of data sources** can be introduced: The **reliable metadata**, such as time for photos, and artist, album, duration and release year for music, can be used to generate fundamental representations of the entire collection. **Similarity** is more suitable as an additional criterion for organization and browsing, and should be offered on demand. In general, similarity can be used to generate a similarity-based view of the entire collection, such as the one offered in *CloudMonster*. Similarity can also be used to conduct an example-based search, or to present content-based suggestions. Besides similarity, content analysis can generate additional information to enhance the browsing flexibility, for example, background color and light status in photos, and gender, genre and mood of music. **User-Generated Data (UGD)** is a special subcategory of additional information, which can efficiently enhance the understanding of semantic relations between items. The UGD used in this part is generated by single users, such as the implicit consumption records (e.g., the popular set in the listening or browsing records) and tags explicitly given by the user. A convenient tagging functionality can encourage the user's contributions, and in turn improve the performance of search.

Multiple representations enable browsing along different criteria. *CloudMonster* offers multiple views for the entire collection and *PhotoMagnets* represents events both in a hierarchical structure and in bubbles. To create these representations, both novel and traditional interface concepts should be combined. For example, *PhotoMagnets* integrates the magnet metaphor in additional to the traditional tree view. In these representations, the commonly used reliable criteria should always be kept in track, for example, by chronologically placing photos in the retrieval results of a time-magnet, or by further grouping songs in the genre view into artist- or album-based subsets. Moreover, the user should be allowed to actively rearrange these representations to create **personalized views**.

Formulating non-specific needs is difficult, and thus efficient assistance should be offered to help the user to **formulate and refine complex query**. A magnet metaphor is proved to be an intuitive and efficient solution. Magnets can be used to execute exploratory tasks beyond a direct search, such as loading items from areas of interest, clustering, comparing and discovering the underlying relations between them. Comparing with *CloudMonster*, the magnet metaphor is enhanced in *PhotoMagnets* with more functionalities, such as null magnet, duplicate magnets and multiple examples, and a new magnetism simulation method. *CloudMonster* offers the magnet-based search within the initial views. The evaluation revealed that the magnet-relevant unstructured activities may destroy the initial structured layout. Therefore, we suggested **a separation of structured and unstructured facilities**. *PhotoMagnets* provides the structured navigation in TreeBrowser and flexible activities in BubbleBrowser. This separation helps to maintain the initial reliable structure, which enables recovery from unstructured activities and distractions. The **semantic coordination** should be improved to facilitate smooth transition between different browsing styles.

The **user interactions** in *PhotoSim* and *MusicSim* are a turn-taking facility and the system cannot response to the user operations in real time. To enhance the interaction continuity and to ensure a seamless user experience, the user interaction and corresponding changes should be synchronized to enable a dynamic response to the user's operations. It is also necessary to offer facilities to help the user making prediction of the results. To improve the **reproducibility**, besides a simple "undo" functionality, saving the browsing history enables recovery from a mistake or switch to a status reached previously. A key frame, such as when a new magnet is added in, can be generated automatically, to reduce the user's explicit effort. However, due to the randomness of magnetism simulation, with the same parameters an exact same temporal visualization cannot be promised. Therefore, a more stable and meaningful positioning of items is necessary. Another concern is the **functionality consistency**, which influences the users' overall impression of a system. For example, in *CloudMonster*, some criteria are provided in initial views but not in color coding schemes or magnet list, which has caused certain degree of confusion among users. Similarly, in *PhotoMagnets* the tagging functionality is only offers in TreeBrowser, but also requested in BubbleBrowser.

5.2 Remarks on the Evaluation Methodology

For *PhotoSim* and *MusicSim*, we conducted an **open discussion** with target users. Their feedback towards the similarity-based interfaces motivated us to combine similarity with reliable metadata and additional information, which led to the other two prototypes of *CloudMonster* and *PhotoMagnets*. Since magnet metaphor is a rather novel concept, we applied an in-depth evaluation in both systems. We conducted an **empirical study** with *CloudMonster* based on a set of pre-defined tasks. We selected those tasks based on several considerations, such as task simplicity, coverage of main functionalities and reasonable completion time. Such constraints may make the observation of user behavior less fruitful. As we realized that interface effectiveness can only be partially reflected by quantitative metrics, such as completion time or accuracy, we tried to collect more subjective data, especially for the users' practical behavior in using such interfaces. During the evaluation of *PhotoMagnets*, besides some concrete tasks, we included two **scenario-oriented tasks** of storytelling and album creation. Relieved from predetermined tasks or specified output, users can explore the system more actively. With a higher degree of motivation, they may discover several advanced usages that may be not foreseen in the original design. Observing the users' free exploration allows an in-depth analysis of user behavior, from which interesting information about the practical usage patterns can be achieved. Concerning the results of empirical and explorative studies, we believe that a combination of both approaches can facilitate a thorough understanding of the system effectiveness and the relevant user behavior.

5.3 Refining the Model of Exploratory Browsing

Through the analysis of system performance and user behavior during the prototype evaluation, we foraged a deeper understanding of *Exploratory Browsing* with personal media collections, which led to a refinement of the original model of *Exploratory Browsing* (cf. Figure 1.3 in Chapter 1). *General Purpose Browsing* is prominent with personal collections, and we discovered its correlation with *Search Browsing* and *Serendipitous Browsing*. We also derived several efficient stimulators for smooth transition between these browsing activities. We refined the initial model of *Exploratory Browsing* in more detail (see Figure 5.1).

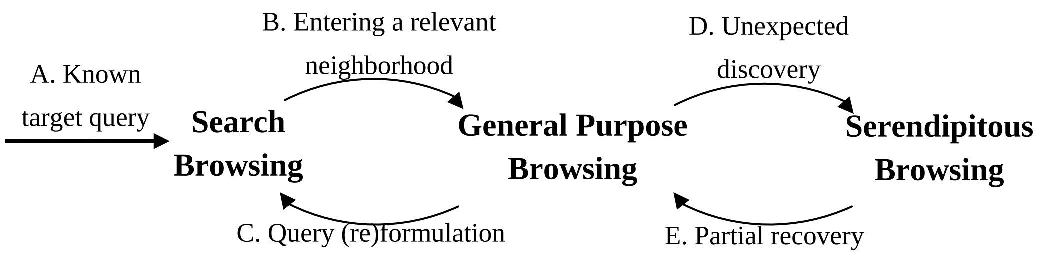


Figure 5.1: The refined model of Exploratory Browsing. The transitional relations between General Purpose Browsing and the other two browsing activities are uncovered.

In this refined model, as a bridge between search and *Serendipitous Browsing*, the importance of *General Purpose Browsing* is obvious. Users may conduct an explicit query or navigate in the entire collection to look for explicit targets (A). Traditional **key-word based search** facilitates looking for a specific target, especially in a rather large graph. A **clear structure of the entire collection** enables structured navigation. With loosely defined goals, users may first discover a relevant neighborhood (B), within which they explore in detail and possibly find some matches. Therefore, enabling **definition of areas of interest** is helpful for *General Purpose Browsing*. During *General Purpose Browsing*, users may gradually formulate their needs and then refine the query (C). Therefore, an **efficient assistance for query formulation**, such as magnets, should be offered. During the exploration, users might get distracted by serendipity (D). Unexpected discovery can be induced for example by presenting items based on their content similarity or semantic relations implied by the user-generated tags. To offer more flexibility, **adjustable parameters** such as the similarity threshold should be allowed. Exploring for a while, users tend to abort the *Serendipitous Browsing* and go back to the normal track (E), for example, by starting a new session or switching back to a previous status. The **anchor point** should always be kept on the original exploration track, which enables the user recovering from distractions.

Because of the users' overall familiarity with their own collections, *Serendipitous Browsing* is not prevalent in personal media collections. To facilitate a detailed investigation of *Serendipitous Browsing* activities, in next part we shifted our focus from personal media collections to online communities, where large numbers of media items are accessible, accompanied with abundant UGD generated among a broader audience. We expected to observe more popular *Serendipitous Browsing* behavior, which would facilitate the investigation of the transitional relations between *Serendipitous Browsing* and the other two browsing activities, and then help to enrich the model of *Exploratory Browsing*.

Part III

Supporting Exploratory Browsing in Online Communities

Part II has presented our exploration to support the *Exploratory Browsing* experience with personal media collections. Chapter 3 introduces *PhotoSim* and *MusicSim*, our first two prototypes built on similarity. The preliminary evaluation confirmed a general appreciation for similarity-based browsing. However, the performance of similarity does not meet the user's high-level perceptions, and they prefer reliable criteria. Therefore, similarity is more suitable as a complementary criterion and should be offered on demand. Chapter 4 presents our progress of coupling similarity with commonly used reliable criteria and additional information, specifically User-Generated Data (UGD). The UGD used in Chapter 4 is user-generated tags and the user's listening and browsing records. We developed *CloudMonster* and *PhotoMagnets* to enhance browsing flexibility with music and photo collections respectively. By observing and analyzing the users' behavior, we discovered that *General Purpose Browsing* is prominent in personal collections. We uncovered its transitional relations with *Search Browsing* and *Serendipitous Browsing*. Because of the users' overall familiarity with their own collections, *Serendipitous Browsing* is found not to be prevalent in personal media collections.

To give us an opportunity to investigate *Serendipitous Browsing* activities, in this part our focus moves from personal collections to online communities. Considering the large quantities of online items and its abundant accompanying metadata, we expected more popular *Serendipitous Browsing* behavior, which can facilitate our investigation of its transitional relations with the other two browsing activities.

The effectiveness of UGD to enhance the browsing flexibility has been confirmed in part II. With respect to abundant UGD generated among large numbers of contributors in online communities (which is called collaborative UGD in this part), we expected that it can help to enhance the exploratory experience with online media collections. We first investigated user behavior and relevant collaborative UGD in such an environment. We conducted a survey of Last.fm[13], our main experimental platform in this part. The results revealed the existence of the *Exploratory Browsing* behavior in online communities. It also indicated that collaborative UGD can improve the users' browsing experience in such an environment. Building upon these insights, we developed four prototypes by integrating collaborative UGD in the browsing interfaces.

The evaluation of these prototypes confirmed the effectiveness of collaborative UGD in supporting *Exploratory Browsing* in online communities. The analysis of the users' behavior in such interfaces confirmed the prevalence of *Serendipitous Browsing* in those platforms. We further discovered its transitional relation with *Search Browsing* and *General Purpose Browsing*, which helped to enrich the model of *Exploratory Browsing*.

---

[13] http://www.last.fm

6

Survey of Online Communities

With the rapid growth of the next-generation Web, it has become a fashion in many websites allowing their users to make contributions by sharing their media collections. The most representative websites are YouTube[14] for video sharing, Flickr[15] for photo sharing and the online music community Last.fm. Their low technical barrier and ease of use have attracted more than hundreds of millions people. Among this huge amount of data, users usually look for known items in a direct search. With unspecified objectives, they can browse the items sorted by categories, popularity or recommendations. Similar to personal collections, we believe that this linear scan offered by current interfaces cannot efficiently support the users' *Exploratory Browsing* behavior. Researchers are looking for new sources and novel representations to enhance the browsing experience in online communities.

Besides sharing own content in those platforms, users can also access others' in turn and hence form very large collections of collaborative UGD. The advent of such data has enormously changed the way people deal with online items. In particular, it is creating new consumption patterns and social interactions, and encouraging users to be more active and creative. The growing popularity of collaborative UGD influences academics and researchers have started to investigate an efficient usage of such UGD to facilitate user activities in online communities. As confirmed in part II, UGD generated by single users illustrates the user's consumption behavior and semantic relations between items on a personal level, and thus facilitates *Exploratory Browsing* in personal collections. Considering the larger quantities and richer diversity of collaborative UGD, we believe that it can efficiently enhance the *Exploratory Browsing* experience with online media collections.

This chapter presents our investigation of user behavior and relevant collaborative UGD in online platforms. We chose Last.fm, an online music community, as our main experimental platform. We conducted a survey of Last.fm and the results revealed that the *Exploratory*

[14] http://www.youtube.com

[15] http://www.flickr.com

Browsing behavior exists in online communities, and that collaborative UGD can help to enhance the users' browsing experience in such an environment. Based on the insights gathered in this survey, we integrated collaborative UGD in the browsing interfaces and developed four prototypes which are discussed in detail in the following chapters.

6.1 Collaborative User-Generated Data

Some existing work addressed statistical analysis of collaborative UGD in online communities. Duarte et al. (2007), Cha et al. (2007) and Cheng et al. (2008) presented in-depth measurements of YouTube videos and their corresponding user profiles. Nazir et al. (2008) statistically explored the temporal and geographic usage distributions of three Facebook[16] applications. Other statistical studies on user behavior in Flickr[17] and Myspace[18] can be found in (Nov et al., 2008), (Prieur et al., 2007) and (Arjan et al., 2008).

Besides the statistical measurements, researchers are looking for an efficient usage of collaborative UGD to facilitate browsing online collections and their relevant metadata. Depending on the author's intentions, we divided collaborative UGD into two categories: implicit UGD and explicit UGD.

Implicit UGD is normally collected through user interaction. For example, the consumption histories are generated implicitly and require no extra effort from users. Through Collaborative Filtering of such UGD (Goldberg et al., 1992; Mehta et al., 2007), the system can understand the users' preferences and usage patterns, derive the underlying relations between users and items, and thus recommend users appropriate products and services. For example, Amazon[19] recommends music and books based on the users' purchase records. Last.fm users receive system suggestions of neighbors with similar taste. Besides improving the quality of system services, implicit UGD can also help users to understand their own and others' consumption behavior. In music communities, consumption history refers mainly to listening history, and a number of applications based on this data have already been proposed. Extra Stats[20] (see Figure 6.1a), LastGraph[21] and Last.fm Spiral[22] extract information of artist or genre from the user's played songs and display their popularity as waves on a timeline. To facilitate comparison of different users' tastes, Last.fm Explorer (Pretzlav, 2008) displays two users'

[16] http://www.facebook.com

[17] http://www.flickr.com

[18] http://www.myspace.com

[19] http://www.amazon.com

[20] http://build.last.fm/item/34

[21] http://lastgraph3.aeracode.org

[22] http://build.last.fm/item/377

listening histories as two respective graphs (see Figure 7.1 in Chapter 7). Besides addressing the temporal characteristic of listening history, other work displays relevant information on a map: Global Matchup Mashup[23] allows comparing the musical taste between a user and a country. musicRambler[24] and LastMusicMap[25] publish relevant images and events of listened artists on a map.

Explicit UGD refers to data actively given by users, such as user profiles, tags, ratings, reviews and comments, etc. Explicit UGD conveys general opinions and knowledge among large numbers of users, from which abundant information can be obtained through certain aggregation. Taking Flickr as an example, Dubinko et al. (2007) presented TagLines, a flow-based animation describing the evolution of the popular tags generated by Flickr users. Ames and Naaman (2007) proposed a photo annotation tool which generates tag suggestions based on relevant Flickr tags. Flickr allows geographically tagging photos by drag-and-drop them on a map, and corresponding geographic locations will be assigned to the photos automatically. With this geographic information, Popescu and Grefenstette (2009) estimated average duration on tourist sights based on a comparison of time difference between photos captured on sight. Combining computer vision and photos' geographic metadata, Yanai and Qiu (2009) selected representative photos for a given term for respective geographic locations, which help to understand how objects, scenes or events under the same concept differ over the world. Based on locations and tags assigned to Flickr photos, Kennedy et al. (2007) extracted representative tags and photos for landmarks in a geographic area.

6.2 Last.fm Survey

Lots of existing work on online communities explores mainly websites such as Flickr and YouTube, and much less is known about how collaborative UGD helps online music communities to make more sense of the musical world. We believe that an investigation on online music communities can lead to a better understanding of people's behavior surrounding music in general and bring valuable insights on how to efficiently utilize the metadata contributed by users to enhance the browsing experience with online media collections.

We chose Last.fm as our main experimental platform. It is one of the most popular online music communities with a large user group and abundant services. The recorded UGD can be accessed through the released API and a series of relevant applications have already been published[26]. However, there is little research systematically investigating user behavior and

[23] http://build.last.fm/item/340

[24] http://www.musicrambler.org

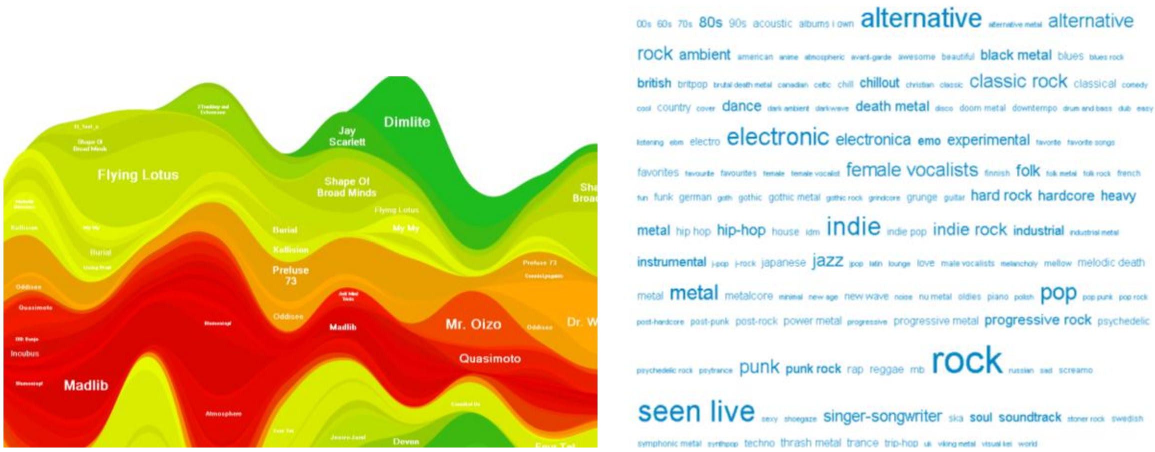
[25] http://www.lastmusicmap.com

[26] http://build.last.fm

relevant UGD in those music communities. To achieve a better understanding of the users' practical needs and the possibility to enhance the browsing experience in those platforms, we conducted a survey of Last.fm (Chen et al., 2010b). We derived substantial insights for the usage of UGD, based on which we developed four prototypes to support *Exploratory Browsing* in such an environment.

6.2.1 User-Generated Data in Last.fm

In Last.fm, each user has a personal profile integrated with a library, playlists, and charts of listened music. Users can listen to music, tag music items and receive recommendations from the system and other users. These activities create abundant UGD, such as user profiles, personal listening histories and tags, which act as the fundament of Last.fm services such as charts and system recommendations.



(a) (b)

Figure 6.1: Visualization of UGD in Last.fm: a) Extra Stats is a temporal visualization of listening history. b) TagClouds[27] display popular user-generated tags in Last.fm.

Listening History

Listening history is recorded automatically when the user listens to Last.fm music. It serves as the statistical basis of Last.fm's main functionalities of charts and system recommendations. Charts are statistical representations of listening history. Personal charts are displayed as lists of music played by the user recently, ordered by play count. Public charts are generated based on an aggregation of personal charts. Based on Collaborative Filtering of the users' listening

[27] http://www.last.fm/charts/toptags - cited 2009/5/25.

histories, the system provides recommendations of similar artists and neighbors with similar taste. Extra Stats is a representative visualization of listening history (see Figure 6.1a): All listened artists are displayed as colored waves on a timeline, with width represents their play counts in each time period.

Tags

Last.fm users can tag music items freely, such as tracks, albums or artists. TagCloud is a well-known visualization for these collaborative tags (see Figure 6.1b), in which tags are normally ordered alphabetically, and the font size represents their popularity. Comparing with tags generated by single users, collaborative tags shown in TagClouds have a higher degree of generality and accuracy. The visualization of TagClouds draws attention to important items and thus reflects the general interests among a broad audience (Viégas and Wattenberg, 2008; Hearst and Rosner, 2008). In Last.fm TagClouds, a keyword-based search can be conducted by selecting a tag as input and the system will return relevant tags and top artists for this tag.

In order to gain insights on an effective usage of collaborative UGD in Last.fm, especially listening histories and tags, we conducted a user study followed by an online survey.

6.2.2 Interview

The interview was conducted in our laboratory. 13 participants were equipped with a PC, a keyboard and a mouse. They were allowed to freely browse the Last.fm website and relevant applications. On average the study lasted about 1 hour per participant. It was recorded on video using the Think-Aloud protocol. All the scores in the questionnaires were rated on a 5-point Linkert-scale where 5 represented the highest score.

At the beginning of the study, the participants filled out a pre-questionnaire about their personal information and general experience with music. Then they joined an interview about their Last.fm experience. They were first asked the main sources of new music and their personal profiles. We then discussed with them about Extra Stats (see Figure 6.1a). Another issue explored was the searching and tagging behavior, and the collaborative tags in Last.fm TagClouds (see Figure 6.1b).

We recruited 13 students in the Last.fm forums, 3 female and 10 male. Their age ranged from 18 to 26 with an average age of 23 years. They were all regular music consumers and rated themselves as experienced Last.fm users (M=4.17, SD=0.72).

Results

Sources of new music: System recommendation of similar artists was the main source of new music for all participants. In addition, 5 out of 13 participants were inspired by other users' suggestions, 2 by browsing neighbors' profiles and only one by searching music of a certain genre.

Personal profile: When asked to describe their musical tastes, most of the participants came up genre-related descriptions and claimed a relatively stable music preference. Although they had no concern about the public nature of their personal profiles, some of them still applied certain strategies to maintain their personal images. For example, one participant intentionally listened to a selected collection to produce representative personal charts.

We discussed with the participants about Extra Stats (see Figure 6.1a), and they were asked about information that could be derived from the visualizations of their own listening history and one of another participant's. A consistent pattern appeared in all participants' listening histories: There were always bursts when the user found new artists and frequently listened to them in a short time period. After a while, these bursts fell into normal flows. The participants thought this tool is interesting and intuitive, and conveys abundant information (detailed comments cf. Chen et al., 2010b): It reflects on the personal consumed music, and facilitates understanding personal listening behavior, recalling relevant social activities, rediscovering forgotten music, and understanding other's musical tastes.

Searching and Tagging: 12 out of 13 participants used the search functionality frequently, and only one found music by browsing the charts of popular artists. Comparing with the most frequently used standard keywords of artist, album and song, tags were much less used for searching (M=2.18, SD=1.08, 5 for daily), and genre was the most often used tag category. All participants felt that user-generated tags are overall too general to guide them finding specific targets.

The participants seldom tagged (M=1.09, SD=0.83, 1 for never) and they mainly tagged music in their own libraries. They though tagging music was difficult. Similar to the description of personal taste and tags used for searching, they usually tagged music with genre information. Some participants also intentionally used distinct tags, such as "listen again" and "Sunday morning", to quickly locate certain music. By observing the Last.fm TagClouds (see Figure 6.1b), all participants noticed that most of the user-generated tags were genre-relevant. They also remarked on the issues of redundancy and errors, especially with genre-related tags.

6.2.3 Online Survey

In order to validate the results of the interview, we conducted an online survey for two months. The questions were consistent with those in the interview, covered mainly the demographic information and general Last.fm experience, specifically source of new music, personal profile, and searching and tagging behavior.

In total we received 228 complete questionnaires, 93 female and 133 male (two gender identifiers were left blank). Their age ranged from 16 to 36 with an average age of 22 years. Most of the participants were students and employees from North America and Europe. They rated themselves to be experienced Last.fm users (M=3.83, SD=0.26).

Results

In general, the results of the online survey were consistent with those received in the interview.

Sources of new music: The main sources of new music were system recommendations (M=3.69, SD=1.30), browsing friends' profiles (M=3.67, SD=1.28), recommendations from friends (M=3.20, SD=1.46), browsing neighbors' profile (M=2.96, SD=1.49) and recommendations from group (M=2.53, SD=1.43).

Personal profile: To describe the personal taste, 173 out of 228 participants proposed genre-related texts. The general attitude toward personal profile being public was rather neutral (M=2.95, SD=1.32).

Concerning the listening behavior, a repetitive listening pattern was reported: They tend to repeatedly listen to certain artists, albums and songs. Consistent with the results of the interview, the visualization of listening history in Extra Stats was perceived as a reflection of listening patterns, which facilitates observing taste changes over time and rediscovering "old" music.

Searching and tagging: The participants looked for music frequently (M=3.99, SD=1.15, 5 for daily), specifically for artist, album and song, and less about social aspects such as groups or events. Different from the interview results, keyword-based search was less popular (M=1.80, SD=1.10), and a significantly larger proportion of time (with an average percentage of 64%) was spent in browsing-like filtering rather than in direct search.

131 out of 228 participants had never tagged and the average tagging frequency was 6.3 times/week (SD=1.22). They mainly tagged music in their own libraries with genre-related texts. Different from the interview, the participants in the online survey considered tagging rather easy (M=2.22, SD=0.09). The top motivations for tagging were easing browsing and searching, assisting personal organization, and conveying personal musical knowledge. The Last.fm TagClouds was commented as useful to derive an overall impression of the most popular tags. The participants also noticed the linguistic problems with user-generated tags.

6.2.4 Implications

Based on the results of this survey, we derived a better understanding of collaborative UGD in online communities, specifically listening histories and tags. We also discovered the existence of *Exploratory Browsing* behavior in those platforms.

Usage of Collaborative UGD

We explored the effectiveness and potential usages of two types of collaborative UGD, listening histories and tags. Listening histories are the most essential UGD in Last.fm, based on which charts and system recommendations are generated. **Graphical representation of listening history** conveys richer information than statistical charts. It offers an intuitive and understandable reflection on the user's consumption records, based on which one's musical taste and listening patterns can be easily read out. It also helps to recall the relevant social events and encourages revisiting “old” music. Although existing visualizations receive positive feedback, they need further improvement. Since users might have a long history, the visualization should offer a good overview. More interactions should be introduced in current static visualizations to improve the browsing flexibility. Most of the tools only target single users and it will be more appealing to offer users an intuitive and convenient way to browse and compare multiple users’ listening histories, which can further improve the understanding of people’s musical tastes and also help to discovery new music.

Users tag music for different reasons and the top motivations for tagging are facilitating browsing and searching, easing personal organization and conveying musical knowledge. Most of the popular collaborative tags in Last.fm are related to genre. **Understanding of genre** is subjective and users may have different understanding of the same genre. There is no standard genre definition and thus less chance for users to be “educated”, which is one of the main factors causing redundancy and errors among genre-related tags. Tag-based search is not popular because tags are perceived as overall too general. They may yield retrieval results with low relevance and thus cannot efficiently help the user locating specific items. **Visualization of collaborative tags** such as TagClouds presents public interests among large numbers of users. We believe that a more efficient organization with less redundancy and errors can convey more general and appropriate knowledge especially with genre, and help to improve the efficiency and accuracy of tagging and searching.

Supporting Exploratory Browsing in Online Communities

Besides personal media collections, *Exploratory Browsing* behavior also exists in online communities. When users have well-defined goals, they locate the target quickly through an explicit search or a linear navigation in certain lists. Users look for items frequently, but they tend to invest more effort on *General Purpose Browsing* than explicit search, especially when they do not intend to look for specific targets, or have not enough knowledge to formulate their requirements. The frequency of *Serendipitous Browsing* with online media collections seems higher than with personal collections, mainly because of the users’ often loosely defined objectives and their overall **lower familiarity** with the large numbers of collections. Discovery of new music is remarkably important for Last.fm users, but currently they mainly find new music among recommendations from the system or other users. Active exploration is less popular and mostly limited as a linear scan of consumption records in others’ profiles or the popular music in the public charts.

Considering the existence of *Exploratory Browsing* activities in online communities and the users’ positive feedback toward UGD-based visualizations, we expect that integrating collaborative UGD in browsing interfaces can enable more flexibility with browsing, encourage users to discover new music more actively and in turn substantially improve their *Exploratory Browsing* experience with online media collections. In the following three chapters, we explored the effective usage of collaborative UGD and developed four prototypes to enhance the *Exploratory Browsing* experience in online communities.

7

Browsing based on Implicit User-Generated Data

In Chapter 6, we grouped collaborative UGD into two categories: implicit UGD and explicit UGD. Our survey confirmed that collaborative UGD helps improving the users' browsing experience in online communities. This chapter presents our exploration to integrate implicit UGC in the interface to enhance the browsing experience.

Consumption history is the most common implicit UGC, which is generated implicitly through the users' interactions and requires no extra effort from the users' perspective. In Last.fm, the users' consumption histories, specifically listening histories play a key role for the generation of system services. Based on an aggregation of all users' listening histories, the system produces public charts for the popular music. Through Collaborative Filtering of listening histories, the system generates recommendations of similar artists and neighbors with similar taste. If the user further browses each neighbor's profile, the similarity of musical taste between these two users is represented as a bar called musical compatibility.

Personal charts list recently most frequently listened music, and in Chapter 4 we utilized this information to produce a popularity view for a music collection in *CloudMonster* (cf. Section 4.1.2). In *CloudMonster* and *PhotoMagnets*, discrete frames in the user's browsing history can be saved, which offers a chance to recover from an operation mistake by switching to a previous status. However, this facility can only reflect respective moments in the user's usage history, and we believe that a more continuous representation can bring an intact image of the user's consumption history. This concept has already been implemented in existing applications (cf. Extra Stats in Chapter 6), in which the consumed items in the entire history is displayed sequentially on a timeline. As the results of our survey revealed, this kind of visualizations helps to discover personal consumption behavior, recall relevant social events and forgotten items, and improve the understanding of others' tastes. Although existing tools receive positive feedback, most of them focus on single users and present static, non-interactive graphics. There is a lack of tools offering intuitive ways to browse, compare and understand multiple users' consumption histories as well as their preferences and tastes.

Beyond the existing visualization targeting single users, we developed *HisFlocks* (Chen et al., 2010a) to support browsing multiple users' consumption histories.

7.1 HisFlocks: Browsing Multiple Consumption Histories

Last.fm Explorer (Pretzlav, 2008) is one of the few works that shows two users' listening histories in two respective graphs (see Figure 7.1). Each graph can be displayed on 3 different levels of tags, artists and tracks. To gather the users' needs and to collect interface concepts, we discussed Last.fm Explorer with 13 regular Last.fm users. The concept of visualizing multiple users' listening histories was overall appealing. The participants appreciated that the history can be displayed on different detail levels. With interactions such as filtering, stacking and brushing, they felt more in control of the visualization.

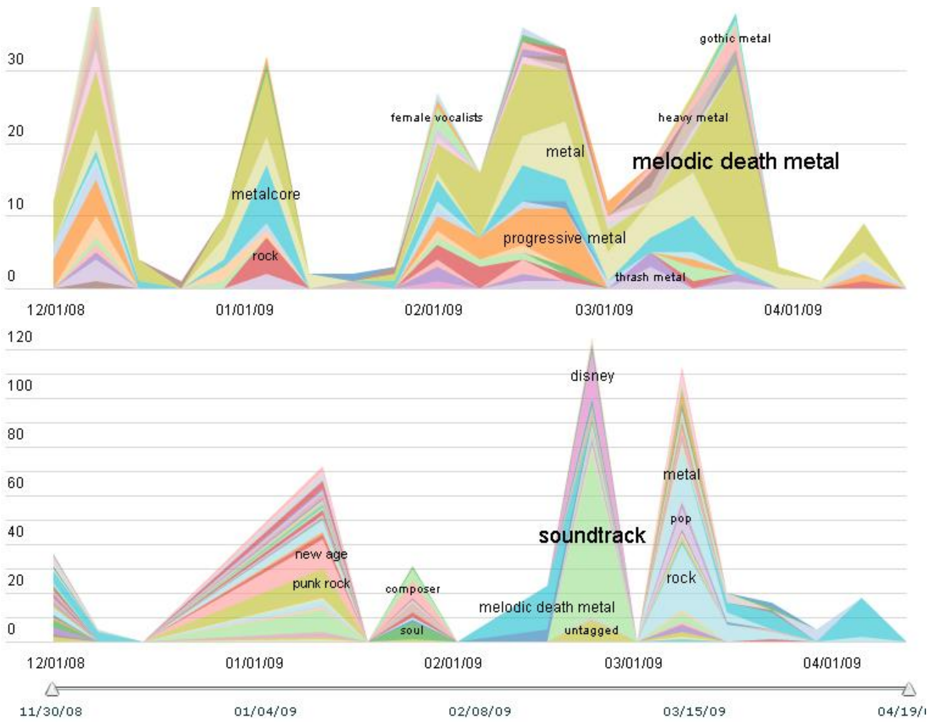


Figure 7.1: Last.fm Explorer (Pretzlav, 2008) displays two users' listening histories in two respective graphs.

The participants' comments also revealed some intrinsic limitations of this tool. The main concern was the difficulty to compare two respective graphs. They also noticed that the timelines were misaligned especially when one user did not listen to music at the beginning or end of the time period. The color coding of the same genre, such as metal and rock, was found inconsistent. These two kinds of inconsistency made it even harder to compare the two respective graphs. Although the history can be displayed on three detail levels, without a semantic zoom, some participants lost the context when switching to another independent level, which forced them to switch back and forth for several times to derive a clear correlation between two levels. Lacking of zoom also leads to poor readability with certain artists and tracks in the stacked graphs. Moreover, only a recent part of history is shown, and the participants wished to see a complete history.

Based on the analysis of existing tools and the discussions with Last.fm users, we gathered the following design concepts to facilitate browsing and comparing multiple users' consumption histories in an interactive visualization:

Easy to compare: Multiple histories should be seamlessly integrated in one interface to enable a convenient comparison between different users.

Keep consistency: The histories should be aligned on a global timeline to ease comparisons within the same time period. The entire history should be displayed and the user should be allowed to select certain time period of interest. Besides time, consistency of other schemes, such as genre color coding, should be also ensured.

Manage complexity: The user's preference and taste can be interpreted on various levels. With the integration of temporal dimension, the complexity steadily increases. Therefore, it is essential to focus on certain aspects, and in the music case, the most promising ones are genre, artist and track.

Provide good overview: To facilitate gaining an overall impression, the visualization should offer a good overview especially for long histories.

Based on these general design considerations, we developed *HisFlocks* (see Figure 7.2) to support browsing and comparison of multiple users' consumption histories, specifically listening histories in the music case. Since genre and artist convey more general information than track, we chose them as the main categories and integrated them on one level. In *HisFlocks*, the entire histories are represented in a series of sequential frames. The time interval between each frame defaults to one week, and it can be adjusted freely in the bottom time slider. Music consumed in each time frame is displayed as corresponding artists on a genre map, where artists from the same genre are grouped into the same genre-cluster, labeled with the genre name. Each artist is color-coded (1 color per compared user) and the size represents its play count. Artists listened to by multiple users are further grouped into highlighted sub-clusters. The user can navigate freely on the map by panning and zooming. In different frames, the same artist always appears at the same position. As Figure 7.2 shows, an overview for all listened artists is offered in the first frame, and by dragging the slider, the user can browse all frames sequentially, in which artists being listened to in each time period are highlighted and others fade out. By observing changes of artists, for example, by their appearance, disappearance and growth in size, the user can observe how one's musical taste changed over time.

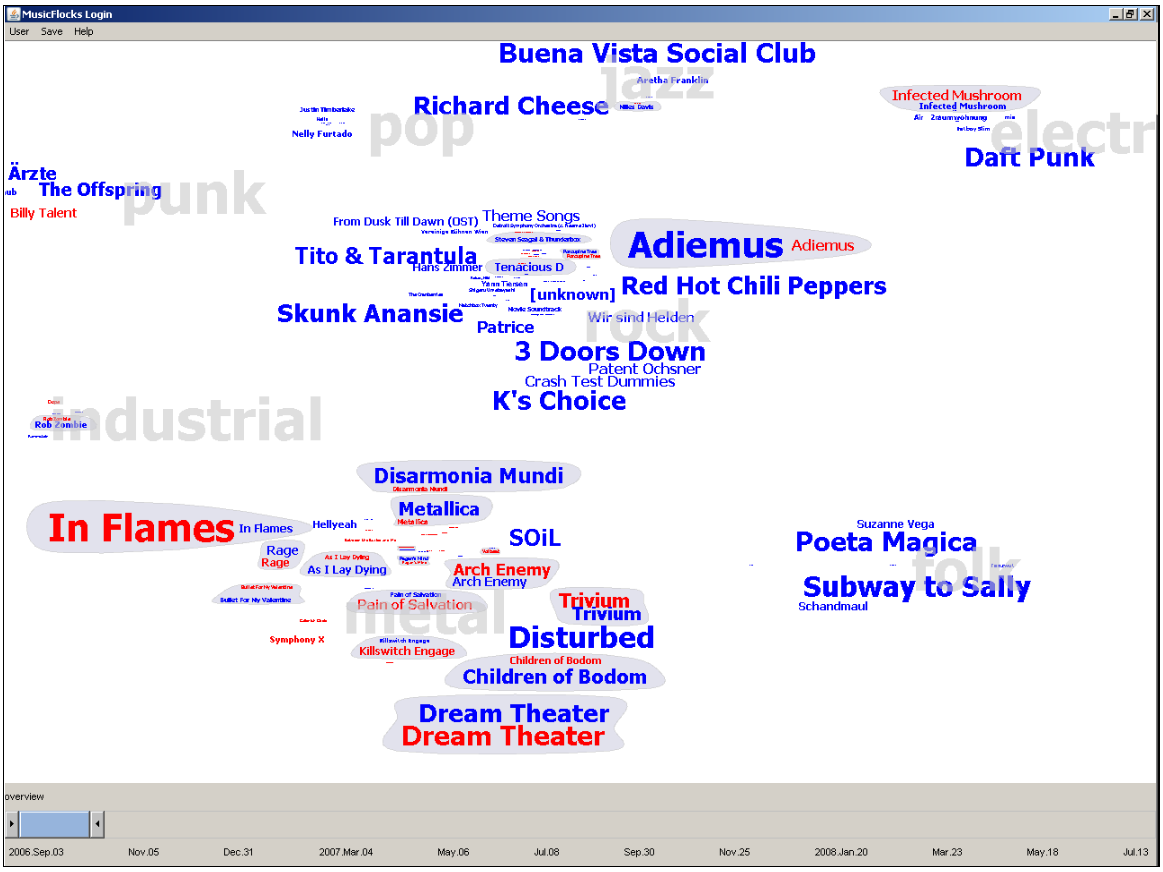


Figure 7.2: The interface of HisFlocks (Chen et al., 2010a), using the same histories displayed in Figure 7.1.

The overview in Figure 7.2 indicates that both users listened to rock and metal a lot, and that they shared many artists in these two genres. They also listened to other genres, such as punk, industrial and electro. The user with the blue color listened to folk, pop and jazz exclusively. Although they shared many artists in common, they listened to few of them in the exact same time period. For example, as Figure 7.3 shows, both users listened to 3 common artists (Killswitch Engage, Children of Bodom, Dream Theater) in the entire timeline, but they only listened to one (Dream Theater) in this time period.

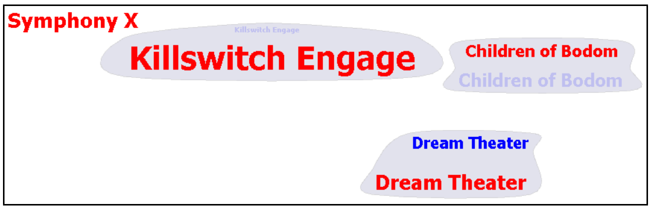


Figure 7.3: Detailed view in one time frame (Chen et al., 2010a).

The two users shared a lot in metal, which is scattered in Figure 7.1 in some sub-genres such as heavy metal, gothic metal and melodic death metal, and thus this information cannot be easily derived at a first glance. To ease browsing and comparison on the general genre level, some aggregation of genre-relevant tags is necessary. In *HisFlocks*, this is achieved based on a semantic analysis. We first applied text analysis to artists' top genre-relevant tags, in order to group them into different genre-clusters: After removing separators such as "_" and "&", we used the Porter algorithm (Porter, 1980) to detect the stem for each tag. Tags with the same stem are clustered in the same cluster. For example, artists from the metal-related genres, such as heavy metal, gothic metal and melodic death metal, are grouped into one metal cluster, thus yielding a more abstract view on the genre level. Locations of genres and artists are determined by their semantic relatedness, which is calculated based on their co-occurrence (see Formula 7.1), a metric widely used in the field of Information Retrieval to determine the semantic relationship between information items (Begelman et al., 2006). In our case, the semantic relatedness between a pair of tags or artists equals to the ratio between the number of resources in which they co-occur and the number of resources in which any of the two items appears. Based on this semantic relatedness, relevant genres or artists are placed near to each other.

$$\text{Co-occurrence}(A,B) = |A \cap B| \,/\, |A \cup B| \tag{7.1}$$

We discussed *HisFlocks* with Last.fm users and received overall positive feedback. Comparing with the musical compatibility bar in Last.fm and the two respective graphs in Last.fm Explorer, *HisFlocks* was perceived as more convenient to browse and compare users' musical tastes in detail. Although in different time frames, the same artist always appears at the same position, the similarity-based positioning of clusters and artists is rather arbitrary. Users cannot anticipate where exactly new artists will appear, similar to the problem of global similarity in *PhotoSim* and *MusicSim* (cf. Section 3.3.1 in Chapter 3). Therefore, a more comprehensible placement will be appealing. Some extra functionality was requested, such as filters.

7.2 Discussion

Consumption histories convey abundant information, from which valuable insights can be read out. We explored an integration of this **implicit UGD** to improve the browsing experience in online communities. We presented *HisFlocks*, a time-based browsing interface, to enhance the understanding of personal tastes, different users' consumption behavior and relevant changes over time. In order to **reduce the complexity**, we focused on the general aspects of genre and artist and integrate them on a semantic genre map, which relieves the user from continuously switching back and forth between different information levels. On the genre map, artists are grouped into corresponding genre-clusters and those listened to by multiple users are further grouped into sub-clusters. Positions of clusters and artists are determined by their semantic relatedness. We applied an **aggregation** for genres to group relevant sub-genres in one general genre, which helped to improve the simplicity and understandability of the visualization. The size of each artist represents its play count, and the color is used for user differentiation. With the **time slider** the user can navigate through the sequential time frames or jump directly to a specific time period. The time interval is adjustable, and an overview of the users' general musical tastes can be achieved by setting the interval of the time slider to all frames.

We compared *HisFlocks* with Last.fm musical compatibility bar and Last.fm Explorer in a **user discussion**. Their feedback towards *HisFlocks* was overall positive. In the current version only two histories are compared and our concept needs to be examined with a larger number of listening histories, which will help to gain additional ideas regarding scalability, understandability and reproducibility. We believe that the basic concept of visualizing multiple consumption histories can be extended to broader scenarios, for example with photos, books or movies.

7.2.1 Usage of Consumption Histories

Users create consumption histories without additional effort but just through their implicit activities, such as listening to music, view photos or purchase books. Our survey of Last.fm and preliminary study with *HisFlocks* revealed that consumption histories help to achieve a deeper understanding of the users' consumption behavior. For single users, consumption history provides **semantic relatedness** of items on a personal level: Items consumed in a batch, such as songs listened to in one playlist, books brought in one purchase, and photos composed in one album, can be considered sharing some common contexts, thus forming a new semantic relatedness of **co-occurrence**. Considering a larger audience, consumption histories convey semantic relatedness between items in a more general way. For example, based on artists' co-occurrence among all users' listening histories, Last.fm recommends similar artists with promising quality. Integrating such UGC in browsing interfaces can help comparing and understanding different users' consumption patterns and preferences, and also facilitate discovering new items through browsing others' records.

7.2.2 Supporting Exploratory Browsing

To facilitate *Search Browsing*, *HisFlocks* allows **temporal navigation** along a timeline, and the user can locate a specific time period with the slider. As discussed in Chapter 5, allowing the user to **define a point of interest** is essential for *General Purpose Browsing*. In *HisFlocks*, the user can adjust the interval for the time slider and freely drag it to locate a specific time span. Prominent items, such as the largest genre group or artists with high play counts, are likely to attract the user's first attention. **Serendipity** can be encouraged by introducing other' histories, from which the user can discover new items and derive taste similarity or difference through a comparison.

This chapter has introduced our exploration of integrating implicit UGD in the browsing interface. In next chapter we investigated the usage of explicit UGD.

8

Browsing based on Explicit User-Generated Data

As discussed in Chapter 7, users can obtain abundant insights from implicit UGD. Although this category of UGD requires no extra effort from users, it has strong dependency on the users' implicit activities and so far only consumption records are widely used. On the other hand, data explicitly generated by large numbers of users conveys general opinions and public interests in a broader range. As the results of our survey in Chapter 6 indicated, explicit UGD such as tags facilitates organization, browsing and searching, and helps gaining additional information and knowledge. In this chapter we integrated explicit UGD in the browsing interfaces to support *Exploratory Browsing* in online communities. We developed *TagClusters* (Chen et al., 2009b) to support semantic browsing of collaborative tags. Besides tags, users explicitly express their detailed opinions, experience and suggestions through ratings, comments, reviews, etc. By combining this type of explicit UGD with the user's personal preferences and needs, we proposed *SARA* to generate personalized package recommendations.

8.1 TagClusters: Semantic Browsing of Collaborative Tags

Our Last.fm survey confirmed that a visualization of collaborative tags, such as TagClouds, conveys general opinions and public interests among a large audience. However, due to the free nature of tagging, there is inevitable redundancy and even errors among these user-generated tags, and this is especially the case with music-relevant tags. Genre is a very important concept for music, which conveys general information that can be expected from a piece of music. Users are likely to use genre to tag music and to describe their music tastes, and thus most of popular tags in Last.fm TagClouds are genre-related (see Figure 6.1b). The understanding of genre is subjective, and there is no standard genre definition and thus no chance for users to be "educated". Redundancy and errors with genre-related tags are prominent in the Last.fm TagClouds. Moreover, the alphabetical order in TagClouds cannot efficiently support a hierarchical navigation and relation exploration.

To give users a chance to learn music-related knowledge and to support semantic browsing in collaborative tags, a semantic-oriented representation of these tags is necessary. Specifically, semantic organization of genre-relevant tags will facilitate semantic browsing and acquisition of genre-relevant knowledge, such as the genre hierarchy and semantic relations between different genres. It can also help to reduce tag redundancy and errors by providing more efficient tag suggestions and in turn improve the efficiency and accuracy of searching and tagging.

8.1.1 Limitations and Improvement of TagClouds

The visualization of TagClouds conveys general interests and popular topics among a broad audience. However, it has several intrinsic problems:

Linguistic problems: Due to the free nature of tagging, two common problems are difficult to avoid from the users' perspective (Wu et al., 2006): Synonymy appears when people use different tags to describe the same item. A term with several different meanings causes ambiguity, which may bring noise into the retrieval results, and consequently reduce the retrieval precision.

High semantic density: If tags are selected only by their usage frequency, there might be a problem of high semantic density (Begelman et al., 2006), which means that very few topics and relevant tags tend to dominate the whole visualization and less important items seem to fade out (Hearst and Rosner, 2008).

Poor semantic understanding: Although the alphabetical arrangement is one important feature of TagClouds, Hassan-Montero and Herrero-Solana (2006) claimed that it poorly supports the understanding of semantic relations between tags. Hearst and Rosner (2008) discovered in their user study that a significant proportion of participants did not even discover the alphabetilization characteristic of TagClouds. They also noticed that users had difficulties to derive the underlying semantic relations among tags, and that spatial distance between tags was misinterpreted as their semantic relatedness. Therefore, TagClouds were considered not an efficient assistance to illustrate the hierarchical structure and semantic relations among tags.

Recent research has investigated improvement of TagClouds in the aspects of aesthetic appearance and semantic understanding. Since several factors influence the effectiveness of TagClouds, such as font size and weight, some systems allow the user to adjust these parameters and one representative application is PubCloud (Kuo et al., 2007). Some algorithms have been introduced to pack tags tighter in the visualization. Kaser and Lemire (2007) proposed

Electronic Design Automation (EDA) to improve the display of TagClouds by avoiding large blank space. Seifert et al (2008) developed several algorithms to display tags in arbitrary convex polygons with dynamically adapting font size.

Lohmann et al. (2009) compared the user performance with different TagClouds variations and discovered that sequential layout with alphabetical sorting is efficient to support locating a specific tag, while a thematic arrangement facilitates finding tags belong to a certain topic. TagOrbitals (Kerr, 2006) display related tags and their summary information in an atom metaphor, in which each primary tag is placed in the center, and other related tags are displayed around. The main problem with this visualization is the text orientation. Hassan-Montero and Herrero-Solana (2006) applied the k-means algorithm to group semantically similar tags. Li et al. (2007) supported browsing large-scale tags based on an analysis of their semantic and hierarchical relations.

Most of the applications discussed above are static and non-interactive visualizations. Moreover, the semantic organization of collaborative tags should be further explored to improve the efficiency of topical exploration and understanding of structure and relations. We explored the implicit hierarchical structure hidden inside the user-generated tags and built *TagClusters* to support semantic understanding of tags, especially genre-related categories.

8.1.2 Enhancing Semantic Understanding of TagClouds

TagClusters are implemented based on Overlapper (Santamaría and Therón, 2008), a visualization tool that highlights the connections and overlapping in data set. In the interface of *TagClusters* (see Figure 8.l), tags are displayed as labeled nodes, and the size represents their popularity. Tags are grouped into clusters and sub-clusters. The color is used to differentiate labels for clusters and tags. As genre-related clusters are prominent and have overlap with each other, they are placed in the center and other less relevant tags and clusters scatter around. Since clusters are represented as transparent colored areas, overlapping clusters (clusters that share one or more common tags) have intersecting, more opaque areas, thus highlighting the overlapping tags. The placement of tags is determined by their semantic relatedness, and relevant tags are placed near to teach other. Therefore, *TagClusters* can be considered a variation of TagClouds in which position conveys semantic meaning. With such a visualization, the user can obtain an intact impression of structure and relation between tags. In Figure 8.1, one can observe that rock is the most popular cluster which is related to several others, such as pop, indie and punk. Also we can find several genres at the bottom right of the figure relating to both metal and rock clusters, such as progressive and goth (see Figures 8.1 and 8.2).

In addition, several interactions are provided, such as panning and zooming, searching, hiding/showing tag labels, replacing nodes and clusters, modifying color and font settings. Infrequently, the force-directed layout may cause an incorrect cluster rendering. In Figure 8.1, if the clusters of goth and folk are placed too closely to each other, there might be a false overlapping between these two clusters. This particular issue is resolved by allowing the user to hover over a node, and the relevant nodes will be highlighted, thereby resolving possible ambiguity. The user can also manually reposition clusters and tags, and relevant changes can be saved automatically.

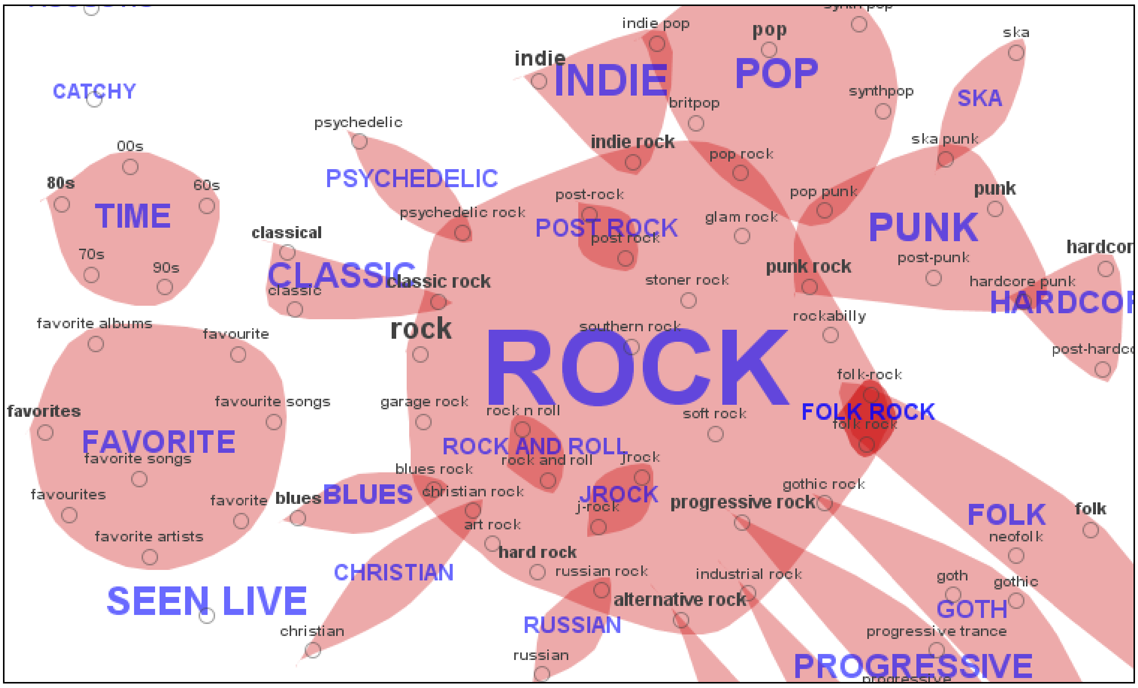


Figure 8.1: The interface of TagClusters (Chen et al., 2009b), with the same tag set in the Last.fm TagClouds (see Figure 6.1b).

Underlying Semantic Analysis

The tag arrangement in *TagClusters* is achieved based on a semantic analysis. The hierarchical structure of tags is derived based on a text analysis, based on which tags are clustered into relevant clusters. The positions of tags are determined by their semantic relatedness.

With an observation of the Last.fm tags, we found that synonymy is the most prominent problem, which is mainly caused by the issue of singular and plural, such as female vocalist and female vocalists, or the language variation between British and American English, such as favorite and favourite. Besides, using different separations between the same words also produces synonymy, for example, post-rock and post rock, or rock and roll and rock n roll. Specifically, the genre-related tags share a common characteristic: The tag in the lower semantic level almost always contains the tag in the higher level and the length of tag is roughly proportional to its semantic level, for example, death metal and brutal death metal. This property helped us to derive the hierarchical structure.

Similar to *HisFlocks* (cf. Chapter 7), we applied a semantic analysis to achieve the arrangement of clusters and tags. After removing separators such as "_" and "&", we apply the Porter algorithm (Porter, 2006) to detect the stem of each tag. Tags with the same stem are clustered into the same cluster or sub-cluster. For example, all the tags containing metal are grouped in the metal cluster and tags such as death metal and brutal death metal are further placed into a death metal sub-cluster (see Figure 8.2). All tags related to gender are clustered in a vocal cluster, and the same with time-related tags, such as 80s and 00s (see left part of Figure 8.1). After the text analysis, most tags can be grouped into relevant clusters properly. This basic approach needs further enhancement to distinguish the literally similar but semantically different tags, such as classic and classic rock. One simple solution is to allow users to manually override the tag structure. With the derived hierarchical structure, tags are clustered into corresponding clusters or sub-clusters. In each cluster, positions of tags are determined by their semantic relatedness, namely their co-occurrence (cf. Formula 7.1 in Chapter 7).

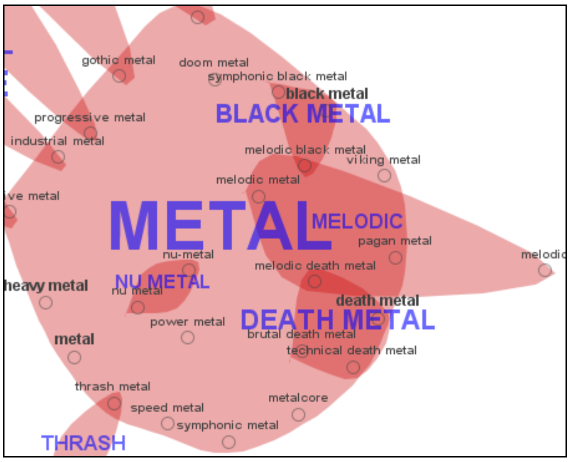


Figure 8.2: Examples of semantic analysis result (Chen et al., 2009b).

With this semantic analysis, tags are grouped into cluster or sub-clusters, and the distances indicate their semantic relatedness. Connected by the overlapping parts, genre-related tags become prominent in the visualization. Other categories such as time- or emotion-related categories are scattered around, due to their less semantic relationship with the genre category.

8.1.3 Evaluation

We conducted a comparative evaluation of TagClouds and *TagClusters*. We were specifically interested in the performance of our tool in supporting the exploration of hierarchical structure and semantic relations among tags.

Design, Settings and Procedure

TagClouds and *TagClusters* were evaluated using a repeated measures within participants factorial design. The independent variable was sysType (TagClouds and *TagClusters*). The order of sysType was counterbalanced between participants to minimize learning effects.

The study was conducted in our laboratory and the participants were equipped with a PC, a keyboard and a mouse. On average the user study lasted about 35 minutes per participant. It was recorded on video using the Think-Aloud protocol. All the scores in the questionnaires were rated on a 11-point Linkert-scale where 10 represented the highest score.

The user study consisted of a pre-questionnaire, an interview and a post-questionnaire. At the beginning of the study, the participants filled out a pre-questionnaire about their demographic information and general experience with collaborative tags. After a brief introduction of TagClouds and *TagClusters*, the participants were asked to execute 6 tasks with both systems, covering the aspects of searching, browsing, comparison and relation understanding. Each task consisted of two similar sub-tasks as following:

Task 1: Locate single tag: Find a tag named German.

Task 2: Sort tags: List the top 5 tags.

Task 3: Compare and filter: List the top 5 genre-related tags.

Task 4: Derive structure: Give the structure of metal-related tags.

Task 5: Detect relation: Is there an overlap between indie and classic?

Task 6: Judge relatedness: Is alternative more relevant to rock or electro?

After completing each task, the participants scored the task easiness and system helpfulness in supporting each task. After completing all tasks, they filled out a post-questionnaire abut their overall impression of both systems.

Participants

We recruited 12 participants at the University of Munich with different majors, 7 German and 5 Chinese, 4 female and 8 male. Their age varied from 24 to 29 with an average age of 27 years. They reported themselves to be familiar with TagClouds ($M=3.58$, $SD=0.99$).

Hypothesis

Initially we had 4 hypotheses: TagClouds will outperform *TagClusters* in Task 1 (H1) and Task 2 (H2), and *TagClusters* will outperform TagClouds in Task 3 (H3) and Tasks 4-6 (H4).

Results

We analyzed the questionnaires and the recorded videos, and the results are shown in Figure 8.3. To locate a tag in task 1, although tags in TagClouds can be located with the alphabetical order, scan for the first character still cost some time. Moreover, 3 out of 12 participants did not notice the alphabetical order in TagClouds, thus spent even more time on scan. A dependent t-test showed that TagClouds cost significantly more time than *TagClusters* ($t_{11}=3.131$, $p<0.05$). The participants claimed that the search functionality in *TagClusters* was especially helpful to locate tags in a rather large graph. Both systems received the same answer precision, but TagClouds was rated lower in the aspects of task easiness and system usefulness. Thus hypothesis (H1) was rejected.

To sort popular tags in task 2, TagClouds performed generally better, and significantly better in the aspects of the time efficiency ($t_{11}=3.92$, $p<0.05$), task easiness ($t_{11}=3.354$, $p<0.01$)

and system helpfulness (t_{11}=3.139, $p<0.05$). Since no significant difference of answer precision was detected, Hypothesis (H2) cannot be proved.

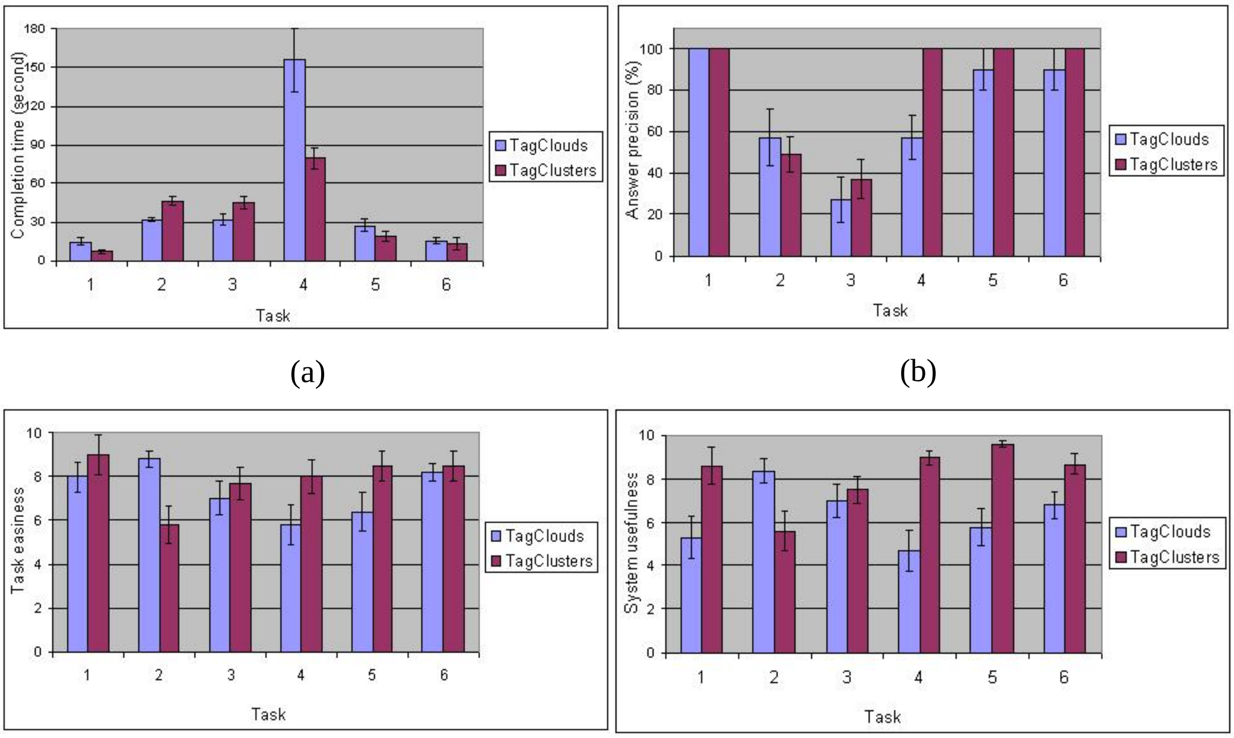


Figure 8.3: The performance of both systems (Chen et al., 2009b) (Error bars represent ± standard error of the mean): a) Completion time (second). b) Answer precision (%). c) Task easiness. d) System usefulness.

In task 3 of tag comparison and filtering, *TagClusters* performed better in the aspects of answer precision, task easiness and system usefulness, but cost more time than TagClouds. Thus Hypothesis (H3) was rejected. The arrangement in TagClouds is compact, which makes a scan of the whole graph easier. To represent the semantic relatedness between tags as their spatial distance, TagClusers need more space and thus create a larger graph. To get a complete image of the entire tag collection, the participants had to keep panning and zooming, and mentally compare and memorize the relevant information, which might slow down the response time and resultanswers with lower precision. Results for task 2 and 3 implied a necessity of making more efficient usage of the *TagClusters* space to create a more compact visualization.

The results of task 4-6 were consistent. To derive hierarchical structure and semantic relations, *TagClusters* achieved better performance in all aspects. A dependent t-test showed that it performed significantly higher in the aspects of system usefulness (task 4: t_{11}=4.872, $p<0.01$; task 5: t_{11}=4.451, $p<0.05$; task 6: t_{11}=2.526, $p<0.05$). *TagClusters* also performed significantly better in task 4 in the aspects of time efficiency (t_{11}=2.752, $p<0.05$) and answer precision (t_{11}=4.077, $p<0.01$). Since semantically similar tags are hierarchically grouped and the overlapping part is highlighted visually, it is easier to judge the relation between tags in *TagClusters*. On the contrary, relevant tags might scatter all over TagClouds, and the participants had to scan all tags to form a mental structure, which cost much more time and led to lower precision. The performance of *TagClusters* was not significant better in all aspects, and thus hypothesis (H4) cannot be verified.

Concerning the overall impression of both systems (see Figure 8.4), *TagClusters* was scored overall higher, and specifically significantly higher in the aspects of enjoyment (t_{11}=2.905, $p<0.05$), understandability (t_{11}=3.446, $p<0.01$) and helpfulness (t_{11}=8.060, $p<0.01$).

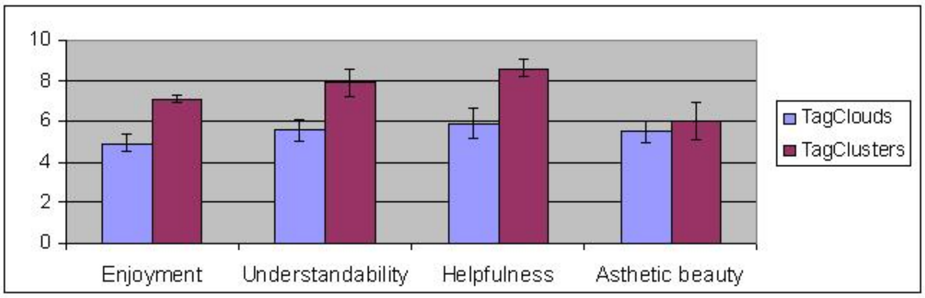


Figure 8.4: Overall impression of both systems (Chen et al., 2009b) (Error bars represent ±standard error of the mean).

Implications

The alphabetical order in TagClouds is an efficient assistance to locate specific tags. However, due to the unfamiliarity with this visualization, some users tend to ignore this feature. Although representing popularity as font size helps to gain an overall impression of the popular tags, others with a smaller font size tend to be ignored in TagClouds. The position also influences the user's visual attention. Some users claimed that the top half part of TagClouds seemed to be more prominent and the bottom half was likely to be ignored. In order to get a compact view by making effective use of the available space, Last.fm TagClouds truncates some long tags and places them into separate lines, which cause certain confusion. The problem of high semantic density is alleviated in *TagClusters* by grouping tags based on their semantic relations. For example, the rock-related tags with small font size still remain relevant in Figure 8.1, since they are clustered into the same cluster with the prominent rock tag.

Sometimes users have problem to locate and compare multiple tags which may scatter all over TagClouds. Without indication of semantic relations, some participants misinterpreted a closer position or similar font size as semantic relatedness. Another problem is that users with less genre knowledge may meet difficulties with relation judgment in TagClouds. This was a prominent problem for most of the Chinese participants. On the contrary, with the assistance of semantic structures in *TagClusters*, they could conduct the same tasks easier. The participants also proposed several aesthetical suggestions, such as stronger highlighting, color coding for different tag categories and a more compact arrangement.

8.1.4 Discussion

We investigated ways to support semantic browsing of collaborative tags. We proposed *TagClusters*, in which tags are grouped into **hierarchical cluster structures**. Font size represents tag popularity and color is used to differentiate labels for tags and clusters. The spatial distance indicates **semantic relatedness** between tags. Several **user interferences** are allowed: The user can adjust the rendering parameters such as color and font settings, and manually adjust tag positions in one cluster or sub-cluster. It would be more useful to allow the user to override the tag structure by repositioning tags across different clusters, which may help to correct some inappropriate clustering results generated by the current semantic analysis. Since *TagClusters* are a semantic aggregation of TagClouds, we conducted a **comparative evaluation** with these two systems, and received overall positive feedback towards *TagClusters*. The results confirmed that our tool has advantages in supporting exploration and understanding of the hierarchical structure and semantic relations among tags. Similar to *HisFlocks* (cf. Chapter 7), the temporal dimension can be integrated to offer a time-based visualization, which may reflect tag changes over time. We believe that the theme-oriented arrangement in *TagClusters* can be transferred to other communities to offer users an opportunity to learn semantic relatedness between topics.

Supporting Exploratory Browsing

Keyword-based search helps to quickly locate certain items in a rather large graph. Displaying the structure within cluster and sub-clusters enables a **hierarchical navigation** through the cluster structure. *TagClusters* is topic-oriented, in which interconnected genre-relevant clusters shape a large flock and become prominent in the visualization. This facilitates gaining an **overview** of relationships among genres. By representing semantic relatedness as spatial distance, *TagClusters* help to derive the semantic closeness between tags. To facilitate **serendipity**, the less relevant clusters, such as time and mood, are remained in the visualization and scatter around the prominent genre-related clusters. They can offer additional options when the user has no specific objectives.

8.2 SARA: Personalized Travel Recommender System

Based on an aggregation of User-Generated Data (UGD) among large numbers of users, more general and reliable information can be extracted, such as the topic popularity and genre-relevant knowledge illustrated in *TagClusters*. However, it always presents the same information to different audience and lacks an identification of the users' personal preferences and needs. If the issue of personalization can be considered in such interfaces, it will efficiently enhance the users' personalized browsing experience.

In online communities, besides tags, users also generate other types of explicit UGD, such as ratings, comments and reviews, which conveys personal experience, opinions and suggestions in detail. In this section we combined this explicit UGD with the user's personal preference and requirements to support generation of personalized package solutions. We specifically investigated this concept in the tourism recommendation domain and developed *SARA* to support personalized trip planning.

8.2.1 Recommender Systems

In daily life, we often rely on recommendations when we need to decide on issues we do not have enough knowledge about. Recommendations with different forms and origins are pervasive from purchasing a book to planning a trip. Because of the abundant promising information provided by large numbers of contributors, Recommender Systems have become one of the main sources for recommendations. Their purpose is to encourage user experience with potentially interesting products and services. Some commercial websites, such as Amazon [28] and Last.fm, have already deployed recommendations to stimulate user consumption. Quality of recommendations largely depends on the underlying algorithm, which has been extensively studied (e.g. Canny, 2002; Huang et al., 2005). However, the ultimate effectiveness of Recommender Systems is influenced by factors that go beyond the algorithm efficiency and accuracy (Cosley et al., 2003).

The interface affects user satisfaction

The only way for users to assess the quality of recommendations is to try them, but users are unlikely to experience without trust. In specific domains, such as tourism, there is no way to try before buy. Therefore, the user's trust in the system is crucial.

Some researchers have already realized the importance of system transparency in stimulating trust. Herlocker et al. (2000) suggested that explanations can make recommendations more understandable, encourage the user's involvement and thus improve the system's performance. Other work further explored the explanation styles in detail (Bilgic and Mooney, 2005; Firan et al., 2009; Montaner et al., 2003).

Swearingen and Sinha (2001, 2002) examined the recommendation quality of several Recommender Systems and discovered that users are willing to provide more input in exchange for more efficient recommendations. Therefore, they suggested allowing users to refine recommendations. Blobworld (Carson et al., 1999) and Fist (Cui and Zhang, 2007) refine query results based on the user's relevance feedback. In Dynamic Query interfaces

[28] http://www.amazon.com

(Ahlberg and Shneiderman, 1994a), the systems dynamically update results according to the user’s actions, such as dragging parameter sliders in FilmFinder (Ahlberg and Shneiderman, 1994b), or adjusting the parameters’ weight by dragging-and-dropping them in a radial graphics in E-Mu jukebox (Vignoli and Pauws, 2005; Herrera et al., 2005). Other work on user interactions in Recommender Systems can be found in (O’Donovan et al., 2008) and (Pauws and van de Wijdeven, 2005).

Recommending package solutions

Besides recommending the individual top N items, some Recommender Systems can automatically generate a package solution (a coherent set of recommendations). Case-based Reasoning (Aamodt and Plaza, 1994) enables constructing a package solution by reusing the solutions built by other users in similar sessions. Although an automatically generated package solution can reduce the user’s effort in decision-making, the full automation will make the exploratory solution seeking process less enjoyable and fruitful. Therefore, we argue that a system supporting users in creating solutions will be more useful in practice. The order of selected items in the final solution is in some cases important. Hansen and Golbeck (2009) introduced the concept of collection-based recommendation, which addresses the correlation between items, such as co-occurrence.

8.2.2 Incremental Recommendation

When recommending more than single items based on given user profile, Recommender Systems mostly either produce a list of the top N items, or automatically generate a package solution. In both cases the underlying recommendation process remains as a black box to users, with which they have no options to interfere. User interactions in these systems are mostly restricted to the adjustment of user preferences, and users cannot dynamically steer the recommendation process.

We believe that full automation can neither practically predict the user’s dynamic preferences nor simulate the complex decision-making behavior of a human. The system’s recommendations should instead be tightly coupled with the user’s interactions to ensure a comprehensive and personalized recommendation process. We proposed the concept of an Incremental Recommender System, in which the recommendation process is decomposed into sequential steps: The system provides recommendations based on adjustable user preferences and the user actively select the item to be included in the final solution, which will in turn influence the system’s recommendations in further steps. In this incremental way, the system’s intelligence and the user’s interaction can be effectively combined to produce a promising personalized final solution. Active control can enable more behavioral flexibility, encourage the users’ participation and thus enhance their exploratory experience.

Three main aspects of an Incremental Recommender System should be considered: User interactions should be allowed to promise a fluid and tight combination of system intelligence and user feedback. Personalization in the recommendation process needs to be ensured. To encourage the user's involvement and satisfaction, the exploratory experience in the recommendation process should be supported.

We believe that this incremental approach can efficiently support the user's decision-making especially for problems requiring inherently sequential solutions, such as creating a shopping route, a music playlist or a travel plan. We investigated this concept in the tourism domain and examined our incremental concept with a Travel Recommender System.

Incremental Travel Recommendation

Recommending tourism products is complicated and multiple factors should be considered, such as the customer's personal characteristics (prior knowledge, personal interest, personality, etc.) and the nature of the trip (e.g., companions, duration and budget) (Fodness and Murray, 1999; Grabler and Zins, 2002). Tourism consumption consists of three phases: prior, during and after the trip (Gretzel et al., 2006). The latter two stages have been extensively investigated (Cheverst et al., 2002; Kroesche et al., 2007; Wohltorf et al., 2005; Kang et al., 2006; Hertzog and Torrens, 2005). Therefore, we focus on the pre-trip phase, specifically the generation of a travel schedule, which requires a reasonable sequential arrangement. Current travel schedules are usually generated fully automatically and the user cannot actively interfere with the scheduling process (Goy and Magro, 2004; Grimnes, 2003; Ricci and Missier, 2004; Sebastian et al., 2008; Schiel et al., 2007). Moreover, the final schedule is mostly represented as a textual agenda, lacking sufficient additional information.

Based on the analysis of exiting applications and our definition of Incremental Recommender System, we believe that an Incremental Travel Recommender System should satisfy the following common characteristics:

Formulate the solution incrementally: To lower the computational cost, encourage the user's involvement and in turn improve the quality of final output, the solution should be generated incrementally: The system proposes recommendations and the user makes choices, and in turn actively ensuring the quality in each step.

Support a high degree of flexibility: Design Galleries (Marks et al., 1997) applied a similar incremental concept, but only produced one single solution. Since people's tourism needs differ according to different motivations and personal interests (Fesenmaier and Jeng, 2002; Gretzel et al., 2006), multi-optional recommendations are necessary, which should be generated based on multiple objective constraints: temporal constraints such as opening time and time spent at the sight; spatial constraints such as geographic distance between sights and time for transportation. Since the user's interests might change dynamically, a high degree of flexibility should be supported by constantly re-adjustable parameters.

Enable an enjoyable experience: Unlike other domains which concern time efficiency and accuracy most, tourism-relevant activities are often not about finding solutions as quickly as possible, but should be an exploratory and enjoyable experience (Gretzel et al., 2006; Hwang et al., 2002). Therefore, a Travel Recommender System should not only suggest efficient recommendations but also support an enjoyable exploration.

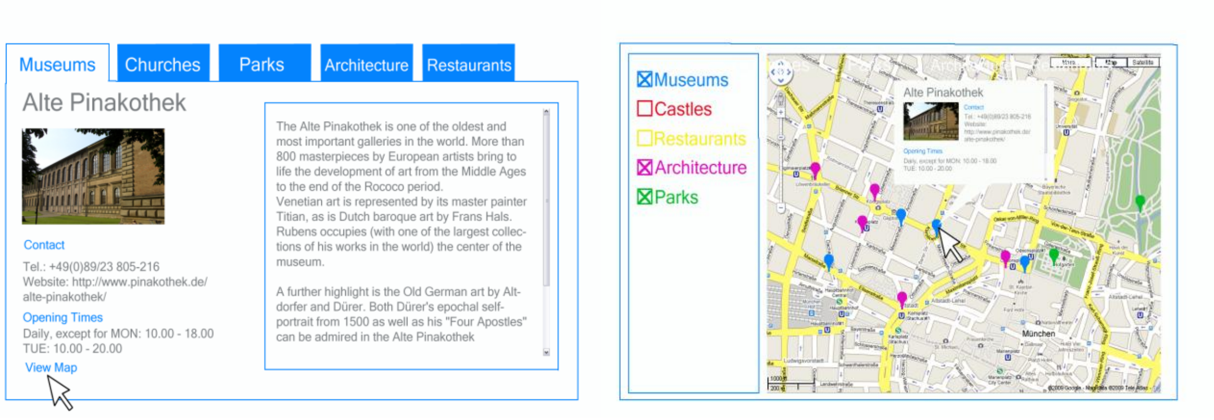
Based on these general design considerations, we presented *SARA* to support creating personalized trip schedule.

8.2.3 Eliciting the Users' Requirements

To gain a better understanding of the users' requirements, we first conducted an expert interview followed by an online survey. We interviewed a travel advisor who has worked for a travel agency in Munich, Germany for more than 20 years. Our questions mainly covered the general consulting workflow and most importantly revealed that there are different customer groups with varying characteristics and needs, for which also the entire recommendation process differs. We decided to narrow our focus to young and generally well-educated customers planning a city sightseeing trip, which is a very common activity with students and young academics.

In order to validate our initial designs, we conducted an online survey for a month. We recruited 100 participants, 46 female and 53 male (one gender identifier was left blank). Their age ranged from 21 to 40, with an average age of 27 years. The participants were mainly students and employees. The questions covered general experience with trip scheduling and some basic concepts of a city sightseeing Recommender System.

The results showed that both personal preferences and objective attributes of sights were important to decide on sights to visit, such as popularity, price, opening time and duration on sight. Public transportation and walking on foot were the most common means of transportation. With respect to a trip planning system, 87% of our participants expressed high willingness to offer detailed personal preferences in exchange for more efficient recommendations.



(a) (b)

Figure 8.5: Two initial interface concepts: a) List-based view: a sight can be shown on a map on demand. b) Map-based view.

Regarding the two interface concepts (see Figure 8.5), both list- and map-based views were appreciated. In the map view, the final plan was represented in a route, and we proposed four different route visualizations (see Figure 8.6): route-only, sight-only, abstract route with sights and precise route with sights. These route concepts were overall appreciated and the precise route with sights was most appealing: route only (M=2.72, SD=1.05), sights only (M=2.84, SD=0.87), abstract route (M=3.45, SD=1.04), precise route (M=4.12, SD=0.96).

(a)

(b)

(c)

(d)

Figure 8.6: Route visualization in a map (Keck, 2009): a) Route only. b) Sights only. c) Abstract route with sights. d) Precise route with sights.

8.2.4 SARA: Stepwise Advanced Route Advisor

Based on the results of the survey, we finalized the design of *SARA* to support creating personalized city sightseeing schedule. *SARA* is implemented based on Google Maps[29] and JSP[30]. Our travel information source is Yahoo Travel[31]. Based on the collaborative UGD, such as travel plans, photos, ratings and reviews, it provided a list of popular sights in one city, with relevant additional information, such as photos, opening time, popularity, rating, address and a textual description.

[29] http://code.google.com/apis/maps

[30] http://java.sun.com/products/jsp

[31] http://travel.yahoo.com

Figure 8.7 shows the interface of *SARA*, which consists of four components (from left to right): a list view, a calendar view, a map view and a column of three setting panels. We explain the scheduling process with the city of Munich, Germany as an example. Initially, the user can set her or his personal preferences for the 5 sight categories in the top right panel, which can always be readjusted dynamically afterwards. The end value of each slider (corresponding to 100%) is the number of most popular sights in the respective city and this number is also shown on the right side of each slider. Initially, all slider values are set to 25%. Each filtered sight is shown as a pin on the map, and the size represents its popularity. Figure 8.7 is an example for a user who prefers three categories.

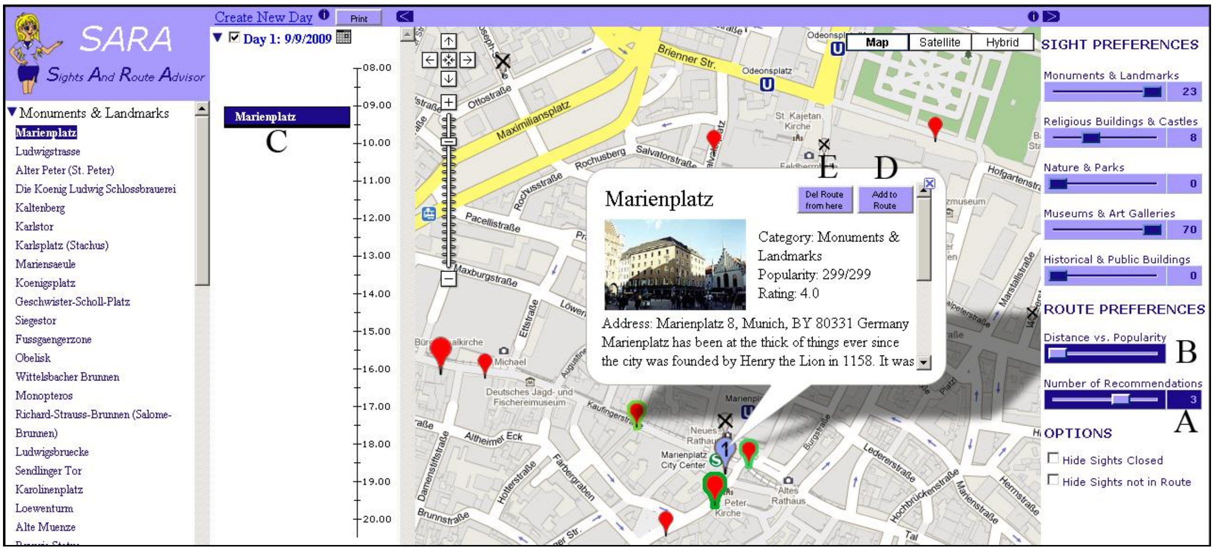


Figure 8.7: The interface of SARA consists of four components (from left to right): A list view, a calendar view, a map view and three setting panels.

With the popular sights filtered based on the personal sight preferences, the user can start the scheduling. The fist step is to set the date for the first day in a pop-up calendar. Before the user adds the first sight to the schedule, she or he can adjust the starting time for this day, which defaults to 8:00am. Comparing this starting time and the opening time of each sight, the sights on the map will be rendered differently: A grey and crossed-out sight is unavailable. Available sights are shown in red and the brightness is a visual reminder about how soon each sight will close (i.e., they “fade” towards closing time). Among the available sights, the system recommends the most popular ones, which are highlighted with a green frame, and the brightness of the frame represents the recommendation rank. The number of recommendations can be changed by a slider (see Figure 8.7-A, the maximum value is currently set as 4 and can be easily enlarged). The user can click on each sight on the map or in the list to see detail information in a pop-up window, and she or he can add this sight to the current plan by clicking button D in Figure 8.7. For example in Figure 8.7, the most popular sight “Marienplatz” is added in the plan. In consequence, it turns blue and becomes numbered as the starting sight for this day. Accordingly, a sight block is added to the calendar view and the block height represents the recommended duration on this sight, which can be changed by dragging the current timeline C. For example, by comparing Figure 8.7 and 8.8, one can observe that the duration for “Marienplatz” was changed from the original 0.5 hour to 4.5 hours. This in turn influences the rendering of some sights and the recommendations for the next sight.

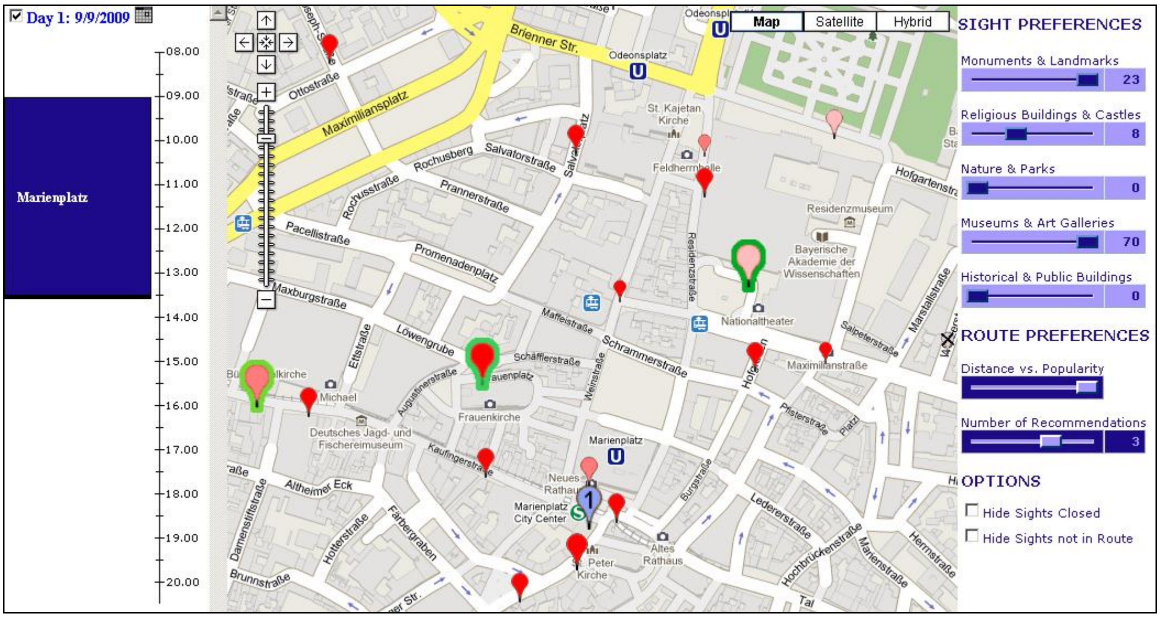


Figure 8.8: System recommendations of next sight based on popularity.

Besides changing the duration on sight, the user can also dynamically steer the system recommendations for the next sight by dragging the slider for route preference (see Figure 8.7-B), which controls the recommendation weight between the sight’s distance and popularity. If more weight is put on distance, the system recommends sights near to the latest sight in the plan (see Figure 8.7). On the contrary, if popularity gets all the weight, sights will be recommended only based on their popularity (see Figure 8.8).

Step by step, a route can be generated which sequentially connects all selected sights. The travelling time between two sights is estimated based on their distance (assuming a pedestrian tourist) and visualized as an extra block between two sights (see Figure 8.9-A). The user can delete sights by clicking button E in Figure 8.7, then the selected sight and all following ones for that day will be deleted from both the calendar and the route. When the user finishes planning the first day, she or he can click “create new day” in the top of the calendar view and continue scheduling the following day. Figure 8.9 shows an example schedule for three days. Each route can be shown/hidden by selecting/deselecting the respective day in the top of the calendar view. The unavailable sights and those not included in any route can also be shown/hidden with the options panel in the bottom left.

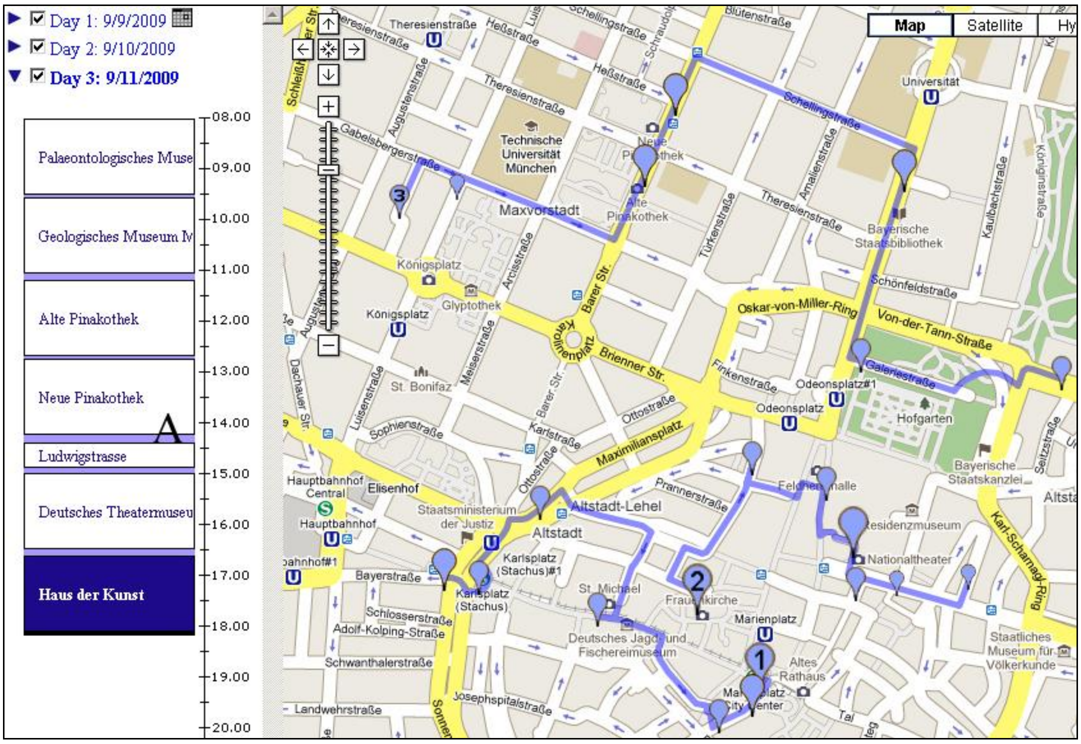


Figure 8.9: An example schedule for three days.

Three Recommendation Modes

SARA supports three recommendation modes, and all of them follow the concept of incremental recommendation and preserve the user's active control. In these modes, the user first sets the initial sight preferences, and then selects one sight as the starting point of the schedule. In the **none recommendation mode**, the user creates the whole plan fully manually without any system recommendations for the next sights. In the **global recommendation mode**, the system recommends the whole schedule as a global solution: After the user sets the end time of the day, the system generates a complete schedule automatically, based on a greedy algorithm that always selects the top ranked recommendation as the next sight. For the generated schedule, the user can change any sight to other alternatives, and the following part of the schedule will automatically update accordingly. The **local recommendation mode** is the one introduced in the previous section, which decomposes the global solution into sequential local recommendations: In each step, the system recommends candidates for the next sight and the user decides on the sight to be included in the schedule. This choice in turn influences the system recommendations for the next sight.

Underlying Recommendation Algorithm

As the system is only responsible for recommendations of the next sight, the complexity of creating a complete schedule is reduced. For the first sight in the plan, the system simply suggests the most popular sights among all available sights. After the user decides on the first sight in the schedule, the system will recommend sights with highest recommendation scores as candidates for the next sight. As Formula 8.1 shows, the recommendation score is calculated based on the sight's popularity and distance value, and the user-defined route preference. To achieve the distance value, the algorithm computes the distance between the current sight and each available sight. This value is normalized by dividing it by the maximum distance between current sight and other available sights. Specifically this value is distracted from 1, in order to get a higher value for shorter distance. The calculation of the popularity value follows a similar method. Both values are multiplied by their weight set by the user, and then are summed up as the recommendation score. The sights with the highest scores will be recommended as the candidates for the next sight.

$$\begin{aligned} Rec\_score_{Sight_i} = & \frac{Sight_i.popularity}{Max\_popularity} * Weight\_popularity + \\ & \frac{1 - Sight_i.distance}{Max\_distance} * (100 - Weight\_popularity) \end{aligned} \tag{8.1}$$

8.2.5 Evaluation

We conducted a user study to evaluate the performance of *SARA*. We were specifically interested in the perceived quality of the system recommendations and the respective performance of the three recommendation modes.

Conditions, Tasks and Study Design

In the study, the participants were required to create a schedule with each of the three recommendation modes. In the none recommendation mode, they created the schedule fully manually without system recommendations for the next sight. In the global recommendation mode, the system created a schedule automatically, and the participants could revise it afterwards by replacing any sight with another available sight. In the local recommendation mode, the system provided suggestions in each step and the participant selected the sight to be included in the schedule.

SARA was evaluated using a repeated measures within participants factorial design. The participants were required to create a schedule with each recommendation mode and the order of mode was counterbalanced between the participants to minimize learning effects. The independent variable was the recommendation mode (none, local, global). The dependant variables were the completion time of each schedule and subjective ratings on a 5-point Likert scale (5 represented the highest score). For each dependant variables we conducted a one-way repeated measures ANOVA and all p values were Bonferroni corrected.

The user study consisted of a pre-questionnaire, an interview and a post-questionnaire. The participants were first asked to fill out a pre-questionnaire with demographic information and their general experience with trip planning. After a brief introduction of *SARA* and its three recommendation modes, the participants played around freely with *SARA*. After they felt familiar with the system, they were asked to create a schedule for a 2-day city sightseeing in Munich with each mode. After completing each schedule, the participants filled out a questionnaire about the performance of each mode. After completing all tasks, they filled out a post-questionnaire concerning the overall impression of *SARA*.

The study was conducted in our laboratory and the participants were equipped with a PC, a keyboard and a mouse. On average the user study lasted about 1 hour per participant. It was recorded on video using the Think-Aloud protocol.

Participants

We recruited 21 participants from our target group (students) with different majors, 11 female and 10 male. Their age ranged from 21 to 27 with an average age of 24 years. All participants were regular users of computers and the Internet, and their average knowledge about Munich is medium.

Hypotheses

We had three hypotheses: A schedule can be created fastest in the local recommendation mode and slowest in the global recommendation mode (H1). Schedules created in the local recommendation mode have the best quality (H2). The local recommendation mode is preferred most (H3).

Results

Plan visualization: Visualizations related to the plan, such as map, calendar and route, were generally appealing. The calendar view with blocks was highly appreciated for obtaining an overview on the time distribution (M=4.48, SD=0.60). During the study, the recommended sight duration was changed 5 times on average and the participants thought it was easy to use (M=4.3, SD=0.86). More flexibility with the calendar view was desired, for example, full control over each sight’s duration instead of only the most recent sight, and changing the sequence of blocks by drag-and-drop. The transportation block was commented as intuitive and the participants expected extra blocks for other activities such as lunch or café. Besides deleting sights from the pop-up window, a more straightforward deleting functionality in the calendar view was requested, for example by right clicking with the mouse.

The map was rated highest among all components (M=4.57, SD=0.60), which was consistent with the results of our survey. With respect to the visualization of sights on the map, although abundant information was integrated in the sight rendering, such as popularity, opening time and recommendation rank, it still received a rather high score (M=4.14, SD=0.79). Especially the visual contrast between available and unavailable sights was commented as intuitive and distinguishable. Route representation was also appreciated (M=4.48, SD=0.81), as it visually illustrates the geographic relation and sequence of the sights in the plan.

Steerable recommendations: The system recommendations for the next sight were generally appreciated (M=4.20, SD=0.81). Although the recommendations were determined by multiple factors, such as popularity, distance, opening time and duration, they were rated as easy to understand (M=4.0, SD=0.67). The user can dynamically influence the system recommendations by adjusting the sight duration or the sight/route preferences, which was appreciated by the participants (M=4.14, SD=0.79). In our study, all participants actually adjusted the preference sliders at least once for each day. The high flexibility was especially practical, as some dynamic planning patterns were observed: Some participants intentionally categorized activities in different days, or even in the same day. For example, one participant scheduled museums exclusively in the morning and other activities in the afternoon. They commented that with the sliders they can arrange the sights more reasonably and conveniently. With another slider the user can adjust the number of recommendations for the next sight. During the study, 16 participants adjusted this slider and 13 out of them even set it to the maximum value. They were highly interested in alternatives and less about their precise ranks.

Three recommendation modes: Each participant was asked to create a plan with each of the three recommendation modes. Based on the collected qualitative and quantitative data, we analyzed the performance of each mode in the aspects of time efficiency, plan quality and usage experience.

In order to compare the **time efficiency**, after creating each plan, the participants were asked to score the statement “Comparing with my normal way of trip planning, I created the plan faster with this recommendation mode”. The average score for the three modes were: none (M=4.19, SD=0.93), local ((M=4.43, SD=0.60) and global (M=3.48, SD=1.29). We found a significant score difference concerning the recommendation mode ($F_{1.345,\ 40}=7.638$, $p<0.05$, with Greenhouse-Geisser correction) and the post-hoc multiple means comparisons showed that the score difference between the local and global modes was significant ($p<0.05$). However, according to the measured completion time, the local mode took longest (none: M=8:07 minutes, SD=0.18; local: M=8:47 minutes, SD=0.22; global: M=8:35 minutes, SD=0.19) and there was no significant time difference among the three modes. The hypothesis H1 cannot be supported. Although the local mode took more time, it was still perceived as fastest. A possible explanation might be that with the assistance of system recommendations in each step, the participants felt easier to make decisions and thus subjectively judged this mode fastest. This also hints at the fact that the ultimate system effectiveness will depend on factors that go beyond efficiency or the quality of algorithms (Cosley et al., 2003).

We also analyzed **the quality of the final schedules**. Travel preferences are very personal, which makes it practically impossible to have plans of one person rated by another. On the other hand, it may be biased to rate the schedules immediately after the creation, since they will always appear optimal at that time. Therefore, we asked participants to score the quality of their own created schedules twice, once in the user study and two weeks later again with randomized order.

The scores in these two rounds are shown in Figure 8.10. We found a significant score difference in the first round concerning the recommendation mode ($F_{1.454,\ 40}=7.012$, $p<0.05$, with Greenhouse-Geisser correction) and the post-hoc multiple means comparisons showed that scores for none and local modes were significant higher than global mode ($p<0.05$). The scores in the second round were generally lower, and a dependent t-test revealed that the scores for none ($t_{20}=3.99$, $p<0.05$) and local modes ($t_{20}=2.65$, $p<0.05$) significantly decreased. However, the mode differences in the second round became smaller and therefore no significant difference was found. The scores in these two rounds revealed a clear bias towards schedules involving more user effort (namely, in none and local modes). We also found a more objective user judgment with the decrease of the novelty. Therefore, we suggest using such an anonymous evaluation to alleviate the users' subjective bias.

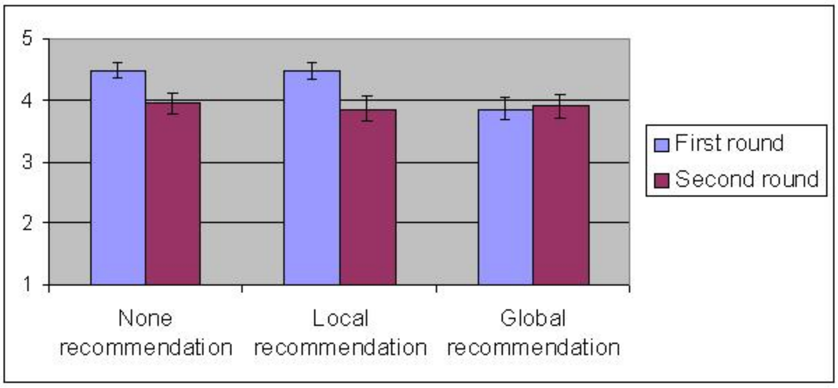


Figure 8.10: Two-round scores for schedule quality (Error bars represent ± standard error of the mean).

In addition to the qualitative user ratings, we defined a quantitative measure for the schedule quality, considering both route and sight quality (see Formula 8.2). With a reasonable route, the user should spend more time on sights and less time on the transportation between sights (see Formula 8.3, cf. Kang et al, 2006). A schedule with higher sight quality should include sights satisfying the user's preferences better. We calculate the sight quality by comparing the similarity between two vectors (see Formula 8.4): V_p is the user's self-reported preference of the five sight categories and V_o are the objective attributes of the plan. For example, if a user rated his sight preference as (70, 80, 50, 80, 80) on a 100-point scale, this is normalized by dividing it by the sum (360), which means that V_p = (19.44, 22.22, 13.89, 22.22, 22.22). Then we count the number of sights in the final schedule for each category (2, 6, 2, 3, 1) and after a similar normalization V_o = (14.29, 42.86, 14.29, 21.43, 7.14). According to Formula 8.4, the sight quality for this schedule is then 0.75. According to this calculation, the average quality of all schedules was rather high: none (M=0.88, SD=0.03), local (M=0.87, SD=0.03) and global (M=0.87, SD=0.03). A statistical analysis revealed no significant difference. Since the schedules generated in the local mode did not achieve the highest quality in both qualitative and quantitative analysis, the hypothesis H2 cannot be supported. However, there is no significant quality difference among the three modes.

$$Schedule\_quality = \frac{Route\_quality + Sight\_quality}{2} \tag{8.2}$$

$$Route\_quality = \frac{\sum_{i=0}^{n} Sight_i.duration}{\sum_{i=0}^{n} Sight_i.duration + \sum_{i=0}^{n-1} TransportationCost(sight_i, sight_{i+1})} \tag{8.3}$$

$$Sight\_quality = Vector\_similarity(V_p, V_o) = \frac{\vec{V}_p \bullet \vec{V}_o}{|\vec{V}_p| \times |\vec{V}_o|} \tag{8.4}$$

To evaluate the **Usage experience**, the participants rated their overall impression of the three recommendation modes (see Figure 8.11). A one-way repeated measures ANOVA showed that the scores were significantly affected by the mode regarding ease of use ($F_{1.658, 33.153}=7.872$, $p<0.01$, with Huynh-Feldt correction), enjoyment ($F_{1.339, 26.779}=10.661$, $p<0.01$, with Greenhouse-Geisser correction) and feeling of control ($F_{2, 40}=16.367$, $p<0.01$). The post-hoc multiple means comparisons revealed both none and local modes had significant differences to the global mode ($p<0.05$). Overall, the local mode was rated highest in all aspects. Since no significant difference was revealed, hypothesis H3 cannot be supported.

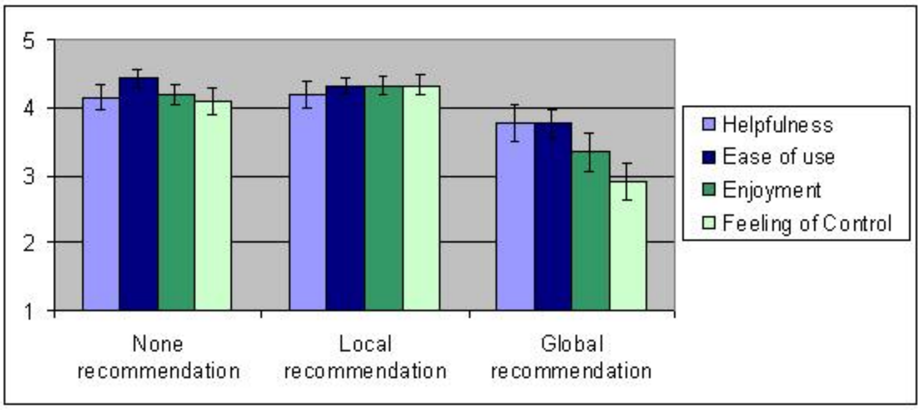
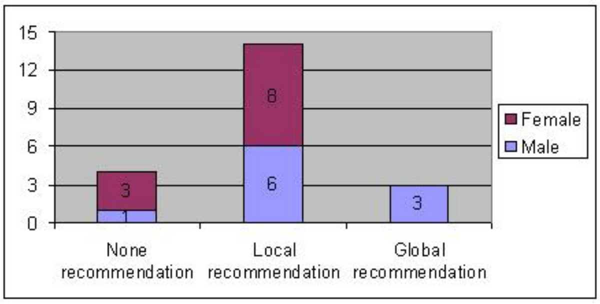
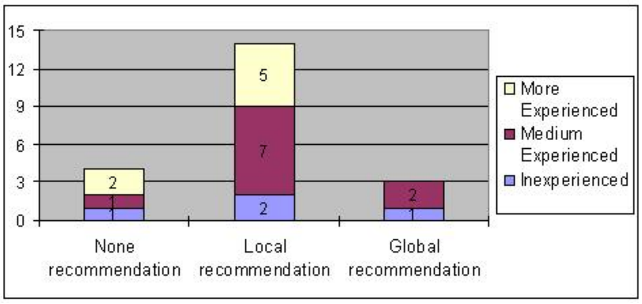


Figure 8.11: Overall impression of the three recommendation modes (Error bars represent ±standard error of the mean).

When asked about the most preferred recommendation mode, 14 out of 21 participants chose the local mode. We further analyzed the relation between mode preferences and gender and trip planning experience (see Figure 8.12). The participants who preferred the global



(a)



(b)

Figure 8.12: (a) Mode preference and relevant users' gender. (b) Mode preference and relevant users' planning experience.

recommendation most were all male or those with lower trip planning experience. On the contrary, the participants who preferred the local recommendation mode were much more experienced. The results indicated a possible correlation between the mode preference and the user's gender or trip planning experience. Although the global recommendation mode was rated lowest in all aspects, it may still be helpful for some user groups.

Implications

Remarks on SARA: It received positive feedback in the evaluation and the concept of Incremental Recommendation was generally appreciated. System recommendations based on the constantly adjustable parameters were highly appreciated since they can reflect the user's dynamic preferences. The local recommendation mode received the most positive feedback and the participants appreciated the active control of each step in this mode. The quality of schedules generated with the three recommendation modes was overall high, which indicates the practical performance of *SARA* in spite of the relatively simple underlying recommendation algorithm. The quality of recommendations can be further improved, for example through Collaborative Filtering (Canny, 2002; Papagelis et al., 2005) or integrating contextual information such as weather and rush hour. Although the global recommendation mode was scored lowest, it was still preferred by some participants. Further analysis revealed that there might be a correlation between mode preference and the user's gender or planning experience.

The interface needs further improvements. Clustering nearby sights in one package would ease browsing and help to alleviate the possible sight occlusion on the map. The slider for route preference can be separated into two independent sliders for sight's distance and popularity respectively, which will offer more flexibility for the user's active control. The route planning functionality should be improved to offer more reasonable and efficient route. Besides, an entire route could be generated based on some key-sights selected by the user. A second round of user study with a totally unfamiliar city will help to discover new usage patterns and probably reveal problems not discovered in the first evaluation. A field study with real tourists will help the investigation of the practical usage of *SARA*.

Incremental Recommendation: By decomposing the whole recommendation process, the user can conveniently steer the recommendation process by making choice in each step with the assistance of the system recommendations, which in turn will influence the system recommendations in subsequent steps. Recommendations generated based on adjustable parameters bring more flexible usages: If the preferences keep unchanged, the recommendations in each step will be generated based on one same scheme, similar to most of existing systems. The user can dynamically change the preferences during the recommendation process. This brings diversity for the recommendations and also helps the user to gradually formulate the detailed and probably dynamic requirements, which is hard to define beforehand. Interactions can ensure the quality of the final solution, promise enjoyable experience and improve the user satisfaction.

Furthermore, different modes of recommendations can be easily realized and the system recommendations can run on different levels. With the recommendation of a global solution, the user can be relieved from processing abundant information. With local incremental recommendation in each step, the user can actively decide on each choice and thus ensure the quality of the final plan. Each mode has its own advantage in satisfying different user groups or the same user under different contexts.

8.2.6 Discussion

We presented the concept of **Incremental Recommender Systems**, in which the system's recommendations and the user's selections are tightly integrated. By decomposing the recommendation process into incremental steps, the user can actively control the recommendation direction through interactions. In contrast to a package solution generated by the system, the user can actively create one. We applied this concept in the tourism domain and presented *SARA* to support creating personalized trip schedule. We conducted a comparative evaluation with the **three recommendation modes** (none, local and global), and the results showed that the local recommendation mode was highly preferred. We also found a possible correlation between mode preference and the user's gender or trip planning experience. We believe that the idea of Incremental Recommendation is especially useful in generating sequential solutions, in which the order of items is important. So far we examined our concept in the tourism domain and we believe that the core concept of incremental recommendation can be extended in other domains, such as music recommendation or general route planning.

SARA is build based on **explicit UGD** generated by users in a tourist community, such as travel plans, ratings, comments and reviews, etc. The abundant information provided by a broad audience helps the user to create personal solution based on others' experience and recommendations. *SARA* offers **different representations of sights**, such as on a map, in a list or shown as blocks in a calendar. The **sight rendering** includes rich information: color for opening time, size for popularity and the frame brightness for recommendation rank. This rendering improves the understanding of the system's recommendations and ease the user's decision-making. To enhance the browsing flexibility and support the behavioral dynamics, **multiple adjustable parameters** are included in the scheduling process, such as opening time, distance and popularity. The user can steer the recommendations by actively deciding the sights to be included in the plan or by adjusting preference sliders or the duration on sight. These adjustable parameters not only facilitate locating candidates by filtering, but also help searching with specific items based on strict constraints. Furthermore, the directness and continuity of **user interaction** is ensured by the concept of Dynamic Query (Ahlberg and Shneiderman, 1994a). The user can observe dynamic changes in the interface while dragging the sight block height or the preference sliders. As sight location in different representations is exclusively determined by precise attributes, such as geographic location or popularity, *SARA* has good **reproducibility**, which means every time when the user adjust the sliders to the same setting, an exact same visualization can be achieved. The scalability especially for sights with close locations needs further improvement, and a possible solution is to group them in one package and browse them in detail through a semantic zoom.

Implications for the Evaluation Approach

We conducted a **comparative user study** with the three different recommendation modes of *SARA*. It enabled an evaluation of the system's overall performance and also helped to derived detailed insights on the effectiveness of different system variations. We **categorized users into different groups** according to their gender and trip planning experience, which helped us to analyze the preferences among different user groups. We then derived a possible correlation between mode preference and the users' characteristics, which also confirmed the advantages of the three modes in different user groups.

During the evaluation of each system variation, we found that **evaluating an interactive interface** is difficult. Typically the system's effectiveness is measured by quantitative metrics, such as completion time and precision. However, user satisfaction is only partially determined by the algorithm accuracy and efficiency. Therefore, quantitative criteria should be combined with the user's qualitative feedback to produce more reasonable and reliable results. For example, the measured time efficiency of the local mode was lowest, but perceived by users as fastest. This conflict between the quantitative and qualitative results however confirmed that the users' active control can efficiently enhance their satisfaction. Specifically, the **exploratory and enjoyable experience** in the interactive process cannot be quantitatively measured. Although we proposed quantitative metrics to evaluate the quality of the final plans, such measures still cannot fully reflect how the final output satisfied the user's dynamically changing interests. Therefore, the user's qualitative judgment, such as the ratings for plan quality and mode preference, should be considered to offer more reliable results. The final plans are always **personalized solutions**, which should be judged only by users themselves. To alleviate the subjective bias, we ran two rounds of qualitative evaluation for the final plan quality. The results indicated clearly that users tend to rate those solutions higher, in which they have invested more effort. The users' judgment is likely to become more objective with the decrease of the novelty. Therefore, we suggest running such an **anonymous evaluation** to gather the users' relative objective feedback. Overall the evaluation of user experience and satisfaction in contrast to the measurement of system efficiency needs a more elaborate exploration. Beyond the laboratory experiment, a field study with real tourists will be helpful for investigating the practical usage of *SARA*.

Reflection on the Model of Exploratory Browsing

Due to the **less confidence** with a large-scale sight collection, users tend to start with *General Purpose Browsing*, even they may have certain sights in mind. Because of the objective constraints such as opening time and geo-distance, they may give up the initial targets and look for other **alternatives**. On the other hand, the *General Purpose Browsing* may be diverted by other sights which share common features with current focus, such as in the same category or locating nearby. If users cannot find any appropriate sights during the *Serendipitous Browsing*, they will reorient with the assistance of certain **anchor points**, such as system recommendations, latest sight added in the plan or a previous representation achieved by resetting the sliders. In general, the *Exploratory Browsing* behavior with *SARA* is overall consistent with the model refined in Chapter 5 (cf. Figure 5.1): *General Purpose Browsing* is prominent, which connects *Search Browsing* and *Serendipitous Browsing*.

Furthermore, we uncovered the transitional relations between *Search Browsing* and *Serendipitous Browsing*. *Serendipitous Browsing* is popular, due to the users' **unfamiliarity** with a large-scale collection of sights. Without explicit targets, some users are completely led by **serendipity**. During the random exploration, their attention tends to be attracted by sights with distinguishable appearance, such as bigger size or darker color. After finding an interesting sight, users may continue exploring the relevant items, such as those locating in the neighborhood or in a same category, thus enter *General Purpose Browsing*. On the way of free exploration, users may recall specific sights and then start to look for them either in the list or on the map. If nothing appropriate is found, users tend to **recover from serendipity** by relocating the previous focus. On the contrary, instead of inspired by serendipity, some users tend to set the most familiar sight as the starting point. However, this sight is not promised to be included in the final plan, since users might dynamically change their mind when facing many options.

Chapter 7 and 8 report the usages of implicit and explicit UGD. The evaluation of the presented prototypes confirmed the effectiveness of collaborative UGD in supporting *Exploratory Browsing* in online communities. Next chapter introduces our exploration to combine these two categories of collaborative UGD to further enhance the browsing flexibility and facilitate deriving additional insights in a broader dimension.

9

Browsing based on a Combination of Collaborative User-Generated Data

In the previous chapters, the development and study of the three prototypes *HisFlocks*, *TagClusters* and *SARA* confirmed that integrating collaborative User-Generated Data (UGD) in browsing interfaces can enhance the *Exploratory Browsing* experience in online communities. Implicit UGD, specifically consumption histories, are especially useful to help understanding personal and others' consumption behavior, and deriving similar patterns among multiple users, such as similar musical tastes illustrated in *HisFlocks*. Explicit UGD conveys general opinions and knowledge among a large audience, such as genre hierarchy represented in *TagClusters*. By taking the users' personal preferences and requirements into account, explicit UGD can be used to make personalized recommendations, such as a personal trip schedule in *SARA*.

Most of existing systems focus on single category of UGD, and the same with these three prototypes. Inherently, implicit and explicit UGD are not independent from each other, and there is an intrinsic connection between them. For example listening histories and tags are connected by the consumed music, thus forming a large-scale of UGD network. We believe that a combination of different types of collaborative UGD can further enlarge the browsing possibilities and help users deriving more insights in a broader range. In this chapter, we integrated both implicit and explicit UGD in a browsing interface to support an exploration of global consumption trends. In general we believe that such an interface should follow several common features:

Aggregation of multi-source raw data: UGD from different sources should be efficiently combined, an aggregation of raw data is necessary to reduce the data complexity and improve the simplicity and understandability of the representation.

Facilitate relational navigation: Considering the interconnection between different categories of information, an easy navigation among them should be supported to ease the exploration of their underlying relations.

Support both geographic and temporal exploration: To facilitate the exploration of large-scale UGD in a global dimension, map is an efficient assistance to ease the comparison between different geographical areas. To enhance the understanding of thematic evolution, the temporal factor should be included to support an observation of changes over time, such as temporal impact of certain items.

Based on these general characteristics, we build *MusicTrends* to support an exploration of worldwide consumption trends.

9.1 MusicTrends: Exploring Worldwide Consumption Trends

MusicTrends is implemented based on the Prefuse toolkit (Heer et al., 2005) for interactive Information Visualization. Three categories of UGD are extracted from the raw data offered by Last.fm API[32]: the users' personal profiles (including user names and origins), their consumed music (represented by the corresponding artists) and relevant tags (specifically the genre-relevant tags). Each category of UGD can be represented in two views: map view (see Figure 9.1) and abstract view (see Figure 9.3a).

9.1.1 Spatial Representation

The interface of *MusicTrends* consists of three components (see Figure 9.1): a main view in the middle, a control panel on the left, and two panels for statistic details and legend on the right. For the category of artist or tag, all items are shown in a list in the control panel. By selecting one artist or tag from this list, the relevant information will be shown in the main view. The user can decide on which view to be displayed by choosing "map" or "abstract" in the control panel. In the map view, countries with top listeners of the selected artist or tag are represented as bubbles. The size of bubbles represents the number of top listeners of this artist or tag in each country, and the color depicts the popularity of the selected artist or tag in each country. For the user category, a list of countries with top listeners is shown in the control panel, and they are also displayed as bubbles in the main view. The size of bubbles represents the number of top listeners in each country and the color describes the similarity between each country's and the global musical tastes.

Abundant insights can be read out from the map view. For example, Figure 9.1 shows the map view for a selected artist "the Beatles". One can observe that a large number of top listeners of this artist come from United States (see the highlighted bubble), and it is not the most popular artist in this country (implied by the brighter color of this bubble). Similar information can be derived with the other two categories of user and tag. With this

[32] http://www.last.fm/api

geographical representation, the worldwide musical characteristics can be clearly illustrated, such as musical similarity between different geographical regions, and the impact of certain artist over the world.

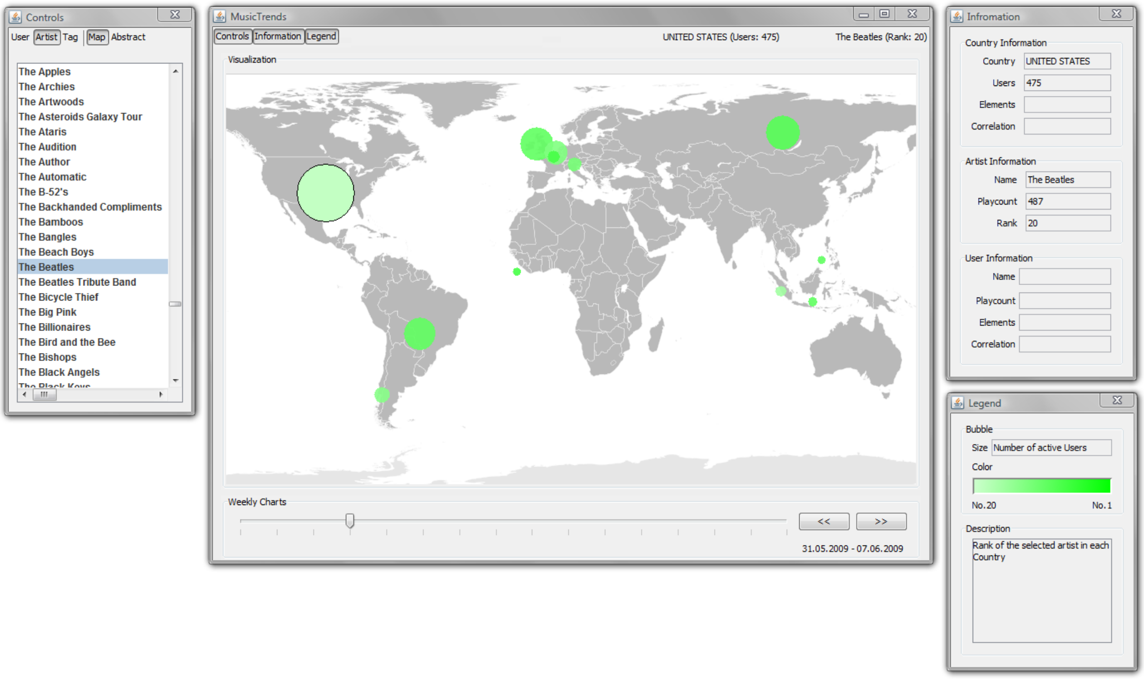


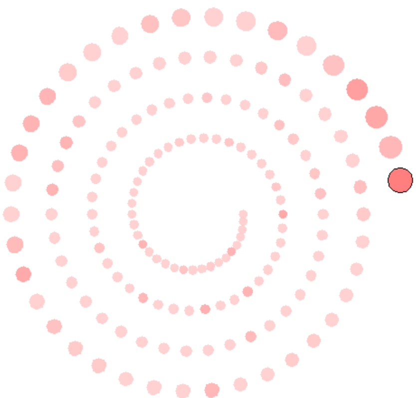
Figure 9.1: The interface of MusicTrends consists of three components: A main view in the middle, a control panel on the left, and two panels for statistical details and legend on the right.

Rendering / Category	Bubble	Bubble size	Bubble color
User	Country with top listeners	Number of top listeners in each country	Similarity between each country's and the global musical tastes
Artist	Country with top listeners of one specific artist	Number of top listeners of this artist in each country	Popularity of this artist in each country
Tag	Country with top listeners of one specific tag	Number of top listeners of this tag in each country	Popularity of this tag in each country

Figure 9.2: Bubble rendering in the map view.

If the user wants to know more details about top listeners, artists or tags in the respective country, she or he can double click on one bubble in the map view, which will leads to the corresponding abstract view. The items in this abstract view are also represented as bubbles, counter-clockwise ordered by their ranks in this country. For the categories of artist and tag, the color and size of bubbles represent the popularity of each artist or tag in this country (see Figure 9.3a). For the user category, the size represents the rank of each user in this country, and the color describes the similarity of musical tastes between each user and this country (see Figure 9.3b). Clicking on the bubble for United States in Figure 9.1 leads to the abstract view for the top artists in this country (see Figure 9.3a). One can observe that “the Beatles” is the 20th top artist in United States (see the highlighted bubble in Figure 9.3a).

(a)



(b)

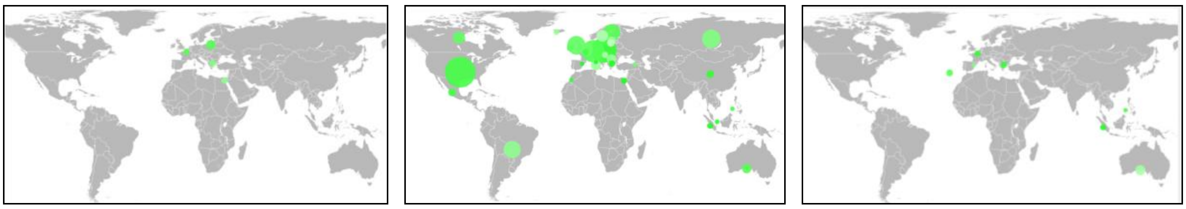
Figure 9.3: Abstract view: a) Top artists in United States. The size and color of bubbles represent the popularity of each artist in this country. b) Top listeners in United States. The size represents the rank of each user and the color describes the similarity of the musical tastes between each user and this country.

Rendering / **Category**	**Bubble**	**Bubble size**	**Bubble color**
User	Top listener in one specific country	Rank of this user in this country	Similarity of musical tastes between this user and this country
Artist	Top artist in one specific country	Popularity of this artist in this country	Popularity of this artist in each country
Tag	Top tag in one specific country	Popularity of this tag in this country	Popularity of this tag in each country

Figure 9.4: Bubble rendering in the abstract view.

9.1.2 Temporal Representation

Besides the geographic representation, *MusicTrends* integrates the time factor to support a temporal exploration, which helps to derive additional time-relevant insights, such as the temporal impact of certain artists or events. The user can select a specific week with the time slider at the bottom of the main view, and the corresponding representation for this week is displayed accordingly. By dragging the slider, the user can observe the weekly changes in the main view. One representative example of such temporal insights is the death of Michael Jackson: Originally he had some regular European listeners (see Figure 9.5a). His sudden death brought a tremendous global impact and larger numbers of users extensively listened to his music in the following weeks (see Figure 9.5b). This impact decreased after his funeral and fell into a stable status (see Figure 9.5c). Comparing with Figure 9.5c, he still gained more listeners spreading over more countries. Some culture-relevant insights can be also derived in this case. For example, one can observe that this artist was overall much more influential in western countries than in Asia.



(a) (b) (c)

Figure 9.5: Sequential frames in MusicTrends illustrated the global impact of the death of Michael Jackson (Schmidt, 2009): a) The week before his death. b) The week of his death. c) The week after his funeral.

Bubble Rendering

The bubble rendering schemes in the map and abstract view are shown in Figure 9.2 and 9.4. Last.fm records top artists for each tag and each country, and top listeners of each artist. Combining these three types of information with the users' origins, lists of top listeners in each country and of each artist and tag can be achieved. Since there is no record for top tags in each country, we define the popularity of a tag in one country as its similarity with the country's taste. We apply Formula 9.1 to calculate the musical similarity between a tag and a country, a country and the world, and a user and a country. The similarity, for example between a user A and his country B, is determined by the rank differences of common artists in their charts (a list of consumed artists). In Formula 9.1, n is the number of common artists, RA_i and RB_i represent the ranks of a common artist i in both charts respectively.

$$Similarity(A,B) = \sum_{i=0}^{n-1} \frac{1}{|RA_i - RB_i| + 1} \tag{9.1}$$

9.1.3 Evaluation

We conducted a user study to evaluate the effectiveness of *MusicTrends* to facilitate exploring worldwide musical trends from large-scale collaborative UGD. Our focus was the category and quality of insights that could be derived from this tool.

Settings and Procedure

The study was conducted in our laboratory and the participants were equipped with a laptop and a mouse. It lasted on average 45 minutes per participant. It was recorded on video using the Think-Aloud protocol. All scores in the questionnaires were rated on a 5-point Linkert-scale where 5 represented the highest score.

The study consisted of a pre-questionnaire, an interview and a post-questionnaire. The participants first filled out a pre-questionnaire about their demographic information and general experience with Last.fm. After a brief introduction of *MusicTrends*, they played around with it for about 10 minutes. After this trial session, they joined an interview. Instead of executing a set of pre-defined tasks, the participants were encouraged to explore *MusicTrends* freely and enumerate insights derived in this free exploration. There was no requirement for speed or quantity of insights. When the participants had difficulties in finding insights, especially at the beginning of the interview, we provided guidance in two levels: We first asked some heuristic questions, such as "can you get some information from a comparison of different countries?" If the participant still had difficulties, we provided a concrete task, such as "List a country where punk is very popular". The guidance ensured the work flow of the study and helped us to examine the easiness to derive different categories of insights. This free exploration lasted about 25 minutes per participant and ended when they claimed that they could not find any new insights. After the interview the participants filled out a post-questionnaire abut their overall impression of *MusicTrends*.

Participants

We recruited 11 participants (6 female and 5 male): 5 students, 1 research assistant, and 5 employees. Their age ranged from 22 to 36 years with an average value of 29 years. 4 participants reported themselves as experienced Last.fm users, 2 medium and 5 inexperienced.

Results

Since *MusicTrends* aims at supporting flexible and enjoyable browsing experience, we put more focus on the analysis of the users' qualitative feedback and the insights derived in their free exploration.

Overall impression: With respect to the participants' overall impression of the two views, map view received higher scores (see Figure 9.6), and a dependent t-test revealed it was rated significantly higher than abstract view in the aspects of enjoyment (t_{10}=3.31, $p<0.01$), helpfulness (t_{10}=2.89, $p<0.05$) and learnability (t_{10}=2.70, $p<0.05$). The results indicated that users may be more interested in a general level of a large-scale data set than detailed information in a specific area.

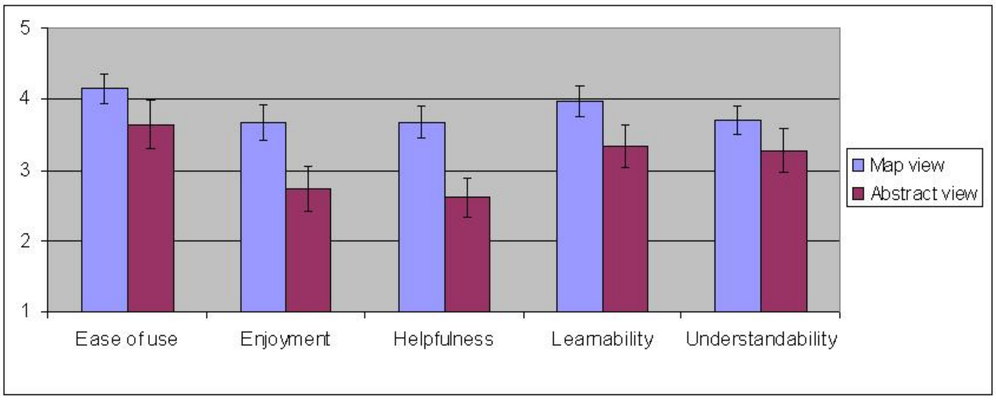


Figure 9.6: Overall impression of the map and abstract views (with error bars displaying the 95% confidence interval).

Inspired by the methodology of insight-based evaluation (Saraiya et al., 2005), we put specific focus on the category and quality of the insights derived by the participants during the free exploration. We also examined the possible correlation between the users' performance and their Last.fm experience.

Insight categorization: In total the participants derived 74 insights. We grouped them into three categories (see Figure 9.7). The first category contained the three UGD types of user, artist and tag. In the geographic category, the location-relevant insights were divided into two subcategories of "country" and "world", depended on whether they were relevant to one single or several countries. There were three sub-categories for the temporal-relevant insights: Those achieved in one time frame were assigned to "point in time". On the contrary, if they were derived based on several frames, they were grouped into either "change" or "constancy", determined by if a change was claimed. Each insight was assigned to one exclusive subcategory of each category. For example, an insight "Pop music is popular in United States" was categorized as tag, country and point in time. For another insight, "the Beatles is always a global popular band", we classified it as artist, world and constancy.

Insight category							
UGC			**Geographic**		**Temporal**		
User	Artist	Tag	Country	World	Point in time	Constancy	Change

Figure 9.7: Three categories used for insight classification.

We classified the achieved insights based on these categories, and the insight distribution is shown in Figure 9.8. Concerning the UGD type, the participants were generally more interested in artists (35) than tags (25) and users (14). They tent to associate artists with the relevant events. For example, participant 9 noticed that a certain band received particular public interests in the month of release of a new album. 9 participants came up 12 insights about Michael Jackson, and 8 out of which were relevant to the global impact of his death (see Figure 9.5). For the geographic category, the participants achieved much more insights in the general global spectrum (52) than in single countries (22), which revealed a general interest on the global level. For example, participant 8 noticed that the users' activeness in many countries regularly drops during holidays. 5 participants came up a similar conclusion that Brazil had the most similar taste to the global area. In the temporal category, 42 out of 74 insights were related to multiple time frames, which highlighted the value of the temporal-based exploration. In detail, there were more change-relevant insights (26) than constancy (16). Specifically, there were more change- relevant implications for artists (16) than tags (7) and users (3). The possible explanation might be the relative stable user preference of certain genres.

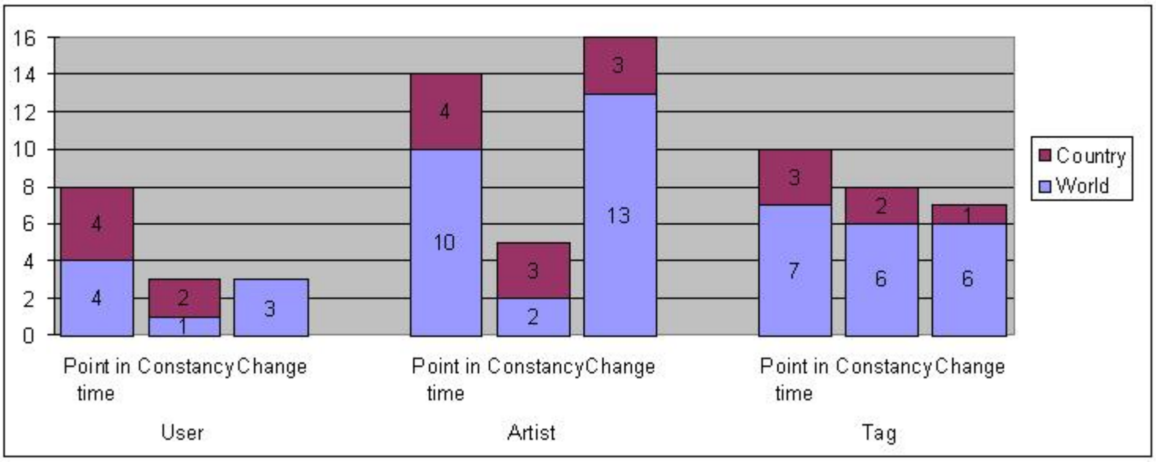


Figure 9.8: Category distribution of all derived insights.

User performance: We measured the insight quality based on its spontaneity and complexity. Each insight was exclusively grouped in one of the three categories of spontaneity, determined by the level of guidance offered by the interviewer during the study: no assistance, heuristic suggestion and specific task. We assigned different weight to them to highlight the value of independency (see Figure 9.9a). We calculated the insight complexity based on the results of insight categorization (see Figure 9.9b). More weight is assigned to the subcategory that is more difficult to achieve, such as world, constancy and change. Taking the two aforementioned insights as example: "Pop music is popular in United States" is categorized as tag, country and point in time, and thus its complexity is 1/3+1/6+1/6≈0.67. "The Beatles is always a globally popular band" is grouped in artist, world and constancy, and its complexity is 1/3+1/3+1/3=1.

Insight spontaneity		
No assistance	Heuristic suggestion	Specific task
1	2/3	1/3

(a)

Insight category							
UGC			**Geographic**		**Temporal**		
User	Artist	Tag	Country	World	Point in time	Constancy	Change
1/3	1/3	1/3	1/6	1/3	1/6	1/3	1/3

(b)

Figure 9.9: Measurement of insight quality: a) Spontaneity. b) Complexity.

We divided the 11 participants into three groups based on their Last.fm experience: experienced (4), medium experience (2) and inexperienced (5). We analyzed the users' performance in each group based on the four criteria: number of insights, time per insight, spontaneity and complexity of insights (see Figure 9.10). The results were not very promising

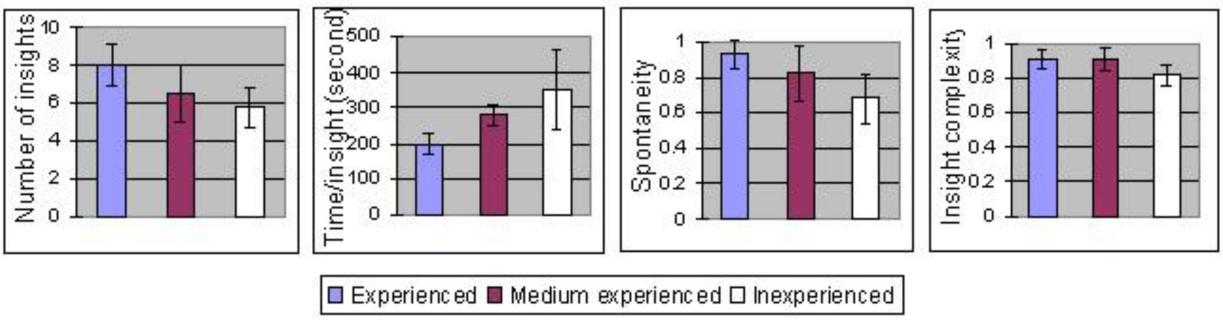


Figure 9.10: User performance in the three groups (Error bars represent ± standard error of the mean): a) Number of insights, b) Time per insight (second), c) Spontaneity of insights, d) Complexity of insights.

due to the limited and unequal numbers of participants in each group. Nevertheless, some interesting findings can still be read out: The quality of the derived insights was overall quite high. There was a possible correlation between the users' performance and their Last.fm experience, and more experienced users tend to derive more insights faster with less external assistance.

Improvement: The bubble was placed in the center of the country's bounding box. In some cases this method did not work accurately. For example, the center shifted from the expected position, if some islands in the oceans were included in the territory of a country. An absolute positioning for example at the location of the capitals can solve this problem, and employing a more detailed map would be necessary. On the current static map, multiple bubbles might be overlapped in a limited space, such as in Europe. Functionality of pan and zoom can alleviate this problem of clutter and ease selecting bubbles with small size.

9.1.4 Discussion

MusicTrends is built based on an **aggregation** of **large-scale collaborative UGD** of users, artists and tags. The underlying interconnection among these three categories of UGD enables users exploring and deriving more insights about musical trends on a global level, which cannot be easily achieved with a linear scan of relevant respective statistical charts. The geographical and temporal dimensions are combined, and the user can navigate on the map or drag a time slider to see the temporal changes over time. This **combination of different representations** helps to expand the browsing dimension, encourages discovering new music and additional insights about the musical evolution over the world. In the map and abstract views, relevant items are shown as bubbles, and the size and color represent different information, such as popularity or musical similarity. Similar to *SARA*, such a geographic representation has good reproducibility with precise positioning. A functionality of semantic zooming is necessary to alleviate possible overlapping in certain areas.

We conducted an **insight-based evaluation** with *MusicTrends*, in which the users were encouraged to look for insights in a free exploration. The overall impression of *MusicTrends* was positive, especially with the map view. To enable a detailed analysis of derived insights, we first group them into 3 categories of UGD type (user, artist and tag), geographic (country and world) and temporal (point of time, constancy and change). This **insight classification** revealed that *MusicTrends* enables gaining different kinds of insights: Larger numbers of artist- and tag-relevant insights indicated the users' higher interests in music-related information. There were more global insights than those relevant to single countries, which revealed a general interest on the global level. Users tent to achieve more insights by browsing multiple time frames, which highlighted the value of the temporal dimension. Specifically, more than half of them were change-relevant insights, and most of them related to artists. One possible explanation is that users have relative stable preference of genres than artists. Besides the insight classification, we analyzed the users' performance based on four criteria: number of insights, time per insight, **insight quality** (spontaneity and complexity). This analysis indicated a possible correlation between the users' performance and their Last.fm experience.

Considering the overall large quantities of collaborative UGD in online communities, we believe that an efficient combination of different categories of such UGD can help users to gain a complete image of the global evolution of public interests and consumption patterns in those platforms.

Reflection on the Model of Exploratory Browsing

Due to the **unfamiliarity** with the large UGD collection of users, artists and tags, some users tend to start with familiar items, such as certain artists or specific geographic areas. Different from *SARA* which involves a concrete objective of creating a trip plan, users explore more freely in *MusicTrends*, without specific tasks or explicit output. Therefore, **aimless exploration** is prevalent, during which users are likely to be attracted by bubbles with bigger size or darker color. Their exploration is easily diverted by **serendipity**, such as noticeable temporal changes and the interconnected relations between the three categories of UGD. Users tend to investigate interesting items in detail, and thus enter the phase of *General Purpose Browsing*. For example, when noticing certain country or geographic area in the map view, some users drag the time slider to see the relevant changes, or click on the country bubble to explore the details of this country in the abstract view. During the free exploration, users may also recall specific items, areas or time period, which leads to a direct search of these targets either in the list, on the map or along the timeline. For example, one participant in our user study saw Michael Jackson in the artist list, which reminded him of the death of this artist. Then he directly dragged the time slider to June to explore the global impact of this event in detail.

10

Discussion

This part has reported on our exploration to support *Exploratory Browsing* in online communities. Our survey of the users' behavior and relevant collaborative User-Generated Data (UGD) revealed that the *Exploratory Browsing* activities exist in such an environment, and that collaborative UGD can enhance the users' browsing experience in those platforms. By integrating collaborative UGD in the browsing interfaces, we developed four prototypes to support *Exploratory Browsing* in online communities: Utilizing the implicit listening histories, *HisFlocks* supports browsing and comparison of multiple users' tastes. Based on explicitly generated UGD, *TagClusters* improves semantic understanding of collaborative tags, and *SARA* facilitates the generation of personalized package recommendations. By combining both implicit and explicit UGD, *MusicTrends* supports exploring worldwide consumption trends. During the development and evaluation of these prototypes, we derived substantial insights for the usage of collaborative UGD, interface design, evaluation approaches and the model of *Exploratory Browsing*.

10.1 Usage of Collaborative UGD

We categorize collaborative UGD as implicit and explicit UGD. Implicit UGD, such as consumption histories, is recorded implicitly through user interactions and requires no additional user effort. By explicitly generating data such as tags, ratings and reviews, users share their personal opinions, suggestions and knowledge actively. For single users, both categories of UGD provide semantic relatedness of items on a personal level. Among a larger audience, this relatedness is represented in a more general way. Comparing with UGD generated by single users, collaborative UGD conveys general interests and opinions with higher accuracy and broader diversity, which enables dealing with online media collections in new ways.

In Part II, we allowed the user to save discrete frames in the browsing history to enable backtracking. In this part, we explored the usage of **implicit UGD**, specifically the continuous and complete consumption histories. We confirmed that they reflect the users' personal interests and consumption behavior. They also produce a semantic relatedness between items, namely co-occurrence. Items consumed in a batch, such as music in one playlist, or photos in one album, share closer semantic relatedness. This information improves the understanding of semantic relations among items and users, and also supports recommending items sharing similar attributes or users with similar taste. An efficient combination of multiple users' consumption histories, such as in *HisFlocks*, facilitates comparison and understanding of different users' personal preferences and tastes, and also helps discovering new items. This concept can be extended to other consumption records, such as photos, books or movies.

Besides implicit UGD, we also explored the usage of **explicit UGD**. Tags convey general opinions and knowledge among a large audience. We proposed *TagClusters* to facilitate semantic exploration of the hierarchical structure and semantic relations among collaborative tags. Such a topic-oriented arrangement can provide users a chance to learn the general knowledge among large numbers of users, such as the semantic relatedness between topics. To improve personalized browsing experience, the user's personal interests and requirements should be included to support personalized services, such as a personal trip plan in *SARA*. The results of the user study confirmed the importance of personalized active control in enhancing the user's *Exploratory Browsing* experience.

Considering the interconnection between different categories of collaborative UGD, such as tags generated for artists that were listened to by different users, we integrated these three types of UGD in *MusicTrends* to support an exploration of global consumption trends. Such a **combination of collaborative UGD** can efficiently expand the browsing range, enhance the understanding of the underlying relations between different categories of UGD, and encourage discovering additional insights in a broader range.

Because of the large numbers of contributors, the size of a collaborative UGD collection is normally quite large. To reduce the data complexity, efficient **combination and aggregation** of collaborative UGD is necessary. For example, in *HisFlocks*, we integrated both genre and artist information on one level and aggregated the genre-related tags to grouping artists into general genre clusters. In *TagClusters* we applied a semantic analysis to determine the hierarchical structure and semantic relatedness between tags, based on which a semantic arrangement of tags was achieved. To facilitate gaining an overall impression in the global dimension, we integrated three types of UGD in *MusicTrends*. Based on the users' consumption records, we further calculated the taste similarity between a user and his country, and a country and the global area, which is represented as the color of bubbles in both map and abstract views.

In general, collaborative UGD can enlarge the browsing possibilities and allow users to explore online media collections in new ways. Such interfaces also encourage discovering new items, obtain additional insights and acquire knowledge which cannot be easily derived through a linear scan of respective collections.

10.2 Considerations for Interface Design

Novel representations beyond a linear scan can facilitate exploration of the underlying relations between items. *HisFlocks* displays listening histories in sequential time frames, and in each frame the listened artists are shown on a genre map. Comparing with a standard list, this representation facilitates exploration and comparison of multiple users' musical tastes. Different from alphabetization in the popular visualization of TagClouds, *TagClusters* group tags into different semantic clusters, and thus enable a semantic exploration. Employing **multiple representations** for the same data set can further improve the browsing flexibility. For example, in *SARA*, sights are displayed on a map, categorized in a list, and shown as blocks in a schedule.

Map-based visualization is extensively used in the presented four prototypes. *HisFlocks* and *TagClusters* group semantically similar items into topic-oriented clusters, and the spatial distance represents the semantic relatedness between items. Similar to the problem in *PhotoSim* and *MusicSim*, arrangement based on the overall similarity is rather arbitrary and the user cannot expect where exactly one item will appear. Therefore, a more comprehensible placement is necessary. Infrequently, the randomness of the underlying force-directed layout might wrongly place some items or create some false overlaps, and a simple solution is the user's manual override. *SARA* and *MusicTrends* employ precise location information to display items on a map. This geographic representation has good reproducibility, but its scalability needs further improvement, especially for items sharing close locations. A semantic grouping with zooming on demand may be one possible solution. Integrating the **time factor** in the map-based visualization can facilitate gaining a complete image about the evolution of public interests and consumption patterns in online communities. For example, *HisFlocks* allows observation of changes of users' musical tastes, and *MusicTrends* enables such an exploration on a global level. To support the behavioral dynamics and often changing user requirements, the relevant parameters, such as the time interval, should be adjustable. To facilitate comparison and exploration especially in different time frames, the same item should always be placed in the same position.

Besides a meaningful positioning of items, the browsing efficiency and the system's understandability can be enhanced by integrating additional information in the **item rendering**. Size and color are the two most often used criteria. Size represents item popularity in the presented four prototypes. Color is used in *HisFlocks* for user differentiation, in *TagClusters* to differentiate labels for tags and clusters, in *SARA* the opening time of each sight, and in *MusicTrends* the similarity of musical taste. Additionally in *SARA*, the brightness of item frame indicates the recommendation rank of each sight.

With respect to the large-scale data set, **personalization** is essential to support an efficient browsing based on personal preferences and interests. *HisFlocks* allows the user to adjust the time interval. In *TagClusters* the user can adjust the color and font settings for tags. In both systems the user can manually reposition each item. *SARA* allows the user to steer the representation dynamically by adjusting sliders for sight and route preferences. Besides, a personalized trip plan can be generated through a tight integration of the user's interactions in the scheduling process.

10.3 Remarks on the Evaluation Methodology

To decide on an appropriate evaluation approach, the main characteristics of the system should be considered. In this part, we have applied different evaluation methods for the presented four prototypes.

A **between-system comparison** can be used to evaluate an improvement of existing techniques. Users compared *HisFlocks* with Last.fm musical compatibility bar and Last.fm Explorer in an open discussion, and we received positive feedback towards the concept of tightly integrating respective consumption histories in one graph. *TagClusters* are a semantic aggregation of TagClouds, and thus we conducted a comparative evaluation with these two systems based on a set of pre-defined tasks. The tasks should be objectively selected to cover the main features of each system. To measure the system performance with concrete pre-defined tasks, quantitative metrics such as completion time and answer precision can be used. Besides quantitative data, the users' qualitative feedback can be collected through user ratings. An analysis of both categories of results enables gaining an overall understanding of the advantages and limitations of each system.

An in-depth study is appropriate to evaluate a system which is very different from existing ones. If variations of the system are applicable, a **within-system comparison** can be conducted. For example, we compared the three recommendation modes in *SARA*. This method helps to obtain an overall impression of the prototype based on an analysis of the performance of each variation. **Classifying users into different groups** according to their characteristics enables an analysis of the correlation between the users' performance and their characteristics. For example, in the evaluation of *SARA*, we discovered a possible correlation between mode preference and the user's gender and trip planning experience. This analysis also confirmed the advantage of each mode for different user groups and even the same user under different contexts.

Since user satisfaction is not only determined by system efficiency, a **combination of quantitative and qualitative metrics** helps to achieve a thorough understanding of the system performance. For example, in *SARA* we discovered an inconsistency between measured completion time and the user's ratings for the local recommendation mode. It cost more time to complete a plan but was perceived by users the fastest mode, which confirmed that user interaction can indeed improve user satisfaction.

Evaluating the users' experience in an interactive interface is not easy and the exploratory and enjoyable experience cannot be efficiently measured by quantitative metrics, such as time efficiency or precision. The users' qualitative feedback is thus important, which can be collected through user ratings, an observation of the users' behavior and an open discussion in a post-session. Measuring the **quality of personalized output**, such as the personal trip plan in *SARA*, is also a difficult task. Objective metrics can reflect the quality of the output to some extend, but it cannot precisely measure how the solution satisfies the user's dynamically changing needs. Therefore, the personalized output should be evaluated by the users themselves. In the user study of *SARA*, we asked the users to score the quality of their final plans in two rounds. The score difference revealed that users tend to rate the output higher, in which they have invested more effort. The users' judgment became more subjective with the decrease of the novelty, which indicated the effectiveness of such an anonymous evaluation to alleviate the subjective bias.

For a browsing interface requiring no concrete tasks or specific goals, such as in *MusicTrends*, an **insight-based evaluation** is more appropriate, in which the user is encouraged to look for insights in a free exploration. Consequently, the analysis of the study results should address the insights derived during this exploration. In the study of *MusicTrends*, we put more focus on the analysis of the **categories and quality of the derived insights**. We first clustered them into different categories, and the insight distribution helped to analyze the users' general interests and the system effectiveness in deriving different categories of insights. We evaluated the users' performance based on four metrics: the total number of derived insights, time cost per insight, insights' spontaneity and complexity. Similar to *SARA*, we grouped users into different groups according to their Last.fm experience. Through an analysis of user performance in different user groups, we found a possible correlation between the users' performance and their Last.fm experience.

10.4 Refining the Model of Exploratory Browsing

During the evaluation of the presented four prototypes, we analyzed the users' behavior and confirmed the model of *Exploratory Browsing* refined in Chapter 5 (cf. Figure 5.1). Because of the high diversity of the large-scale data set and the users' generally low familiarity, we found that *Serendipitous Browsing* is more popular in online communities than within personal collections. We further discovered its transitional relation with *Search Browsing* and *General Purpose Browsing*. With the uncovered full range of transitional relations between browsing activities, the model of *Exploratory Browsing* is further enriched. The three browsing activities become interconnected, thus forming an intact model of *Exploratory Browsing* (see Figure 10.1).

The users' intensions influence their browsing behavior. With clear objectives, such as creating a travel plan in *SARA* or judging the similarity between two users' musical tastes in *HisFlocks*, users tend to start with *General Purpose Browsing*. They are likely to explore items sharing similar attributes with the temporal focus, such as sights locating nearby or artists in the same genre. Because of the large quantity and high diversity of collaborative UGD, *Serendipitous Browsing* is more active in online communities. Especially when no specific goal or explicit outcome is involved, such as in *MusicTrends*, the exploration is more likely an aimless wander (H), during which users fully rely on serendipity.

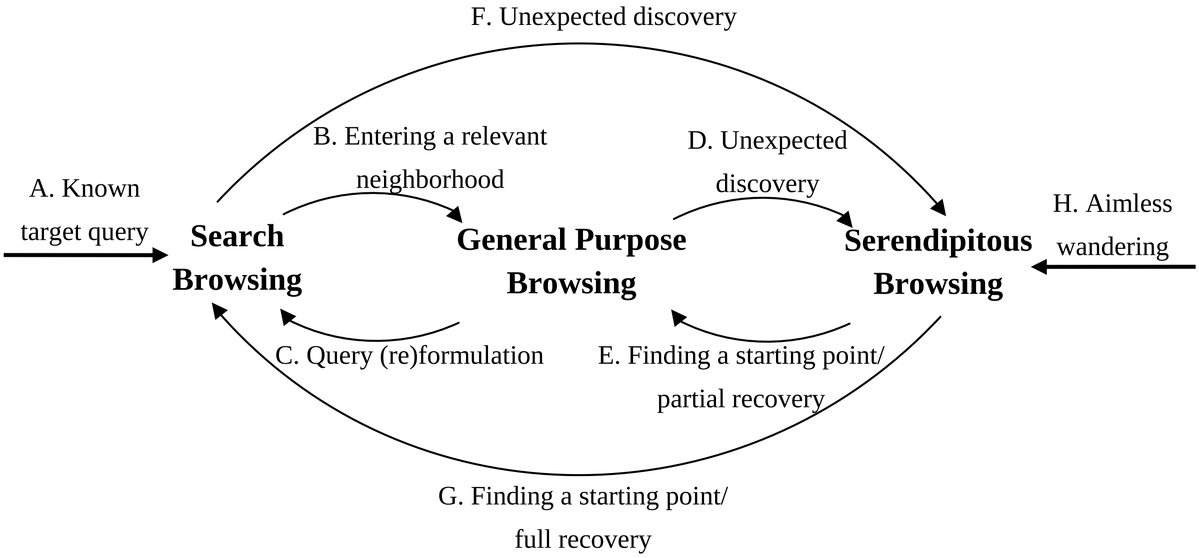


Figure 10.1: The model of Exploratory Browsing. The full range of transitional relations between the three browsing activities is uncovered.

The users' initial objectives might be frequently changed when facing many alternatives. The transition from *Search Browsing* and *General Purpose Browsing* to *Serendipitous Browsing* is mostly encouraged by the **unfamiliarity** introduced by large-scale collaborative UGD (F, D). This information helps deriving diverse findings in a broad range, such as users with similar taste and the global evolution of general interests. A **combination of different visualizations**, such as geographic and temporal representations in *SARA* and *MusicTrends*, facilitate exploring the underlying relations between items and thus also encourage serendipity. During the free exploration, prominent items and areas are more likely to draw the users' attention, such as items with instinct appearance and groups including abundant items. After locating an interesting item, users may explore the relevant area in more detail, thus step in the *General Purpose Browsing* (E). To facilitate defining an area of interest, **adjustable parameters** should be introduced, such as the time interval in *HisFlocks* and the preference sliders in *SARA*. They may also recall certain items and look for it in a *Search Browsing* (G). To quickly locating a specific target, *TagClusters* offers **keyword-based search** and the other three prototypes allows **jumping directly** to a specific item in a list or locating a specific time period on the timeline. If users cannot find any appropriate items or insights, they tend to abort *Serendipitous Browsing*, switch back to the **previous accessed points** or start a new exploration (E, G). The user can drag the time slider in *HisFlocks* and *MusicTrends* to switch to a certain time period. In *SARA* a previous visualization can be reproduced by adding or deleing sight in the plan, or by adjusting the preference sliders.

Part IV

Conclusion and Future Work

This dissertation has presented our exploration to support the *Exploratory Browsing* experience with media collections. Part I introduces the motivation, research context and methodology of this thesis. It also presents our definition and initial model of *Exploratory Browsing*. This introduction is followed by a discussion of related work on interface design for media collections. Part II presents our investigation to support *Exploratory Browsing* in personal media collections. The initial model of *Exploratory Browsing* is refined by uncovering the transitional relations between *General Purpose Browsing* and the other two browsing activities. Part III addresses the usage context in online communities and discusses our exploration to integrate collaborative User-Generated Data (UGD) in the browsing interfaces to support *Exploratory Browsing* in online communities. The model of *Exploratory Browsing* is further enriched by discovering the transitional relations between *Serendipitous Browsing* and the other two browsing activities.

This last part concludes this dissertation. Building upon the explorations in the previous parts, Chapter 11 summarizes the process and approaches that were undertaken in this dissertation. It discusses the main outcomes of this research. Based on the model of *Exploratory Browsing*, a catalogue of general interface characteristics is presented. Also, several general suggestions for designing interfaces for media collections are proposed. Finally Chapter 12 concludes this dissertation with discussions on future work to elaborate the current model of *Exploratory Browsing* and explore its applicability in relevant domains.

11

Summary and Conclusions

This chapter builds on the results of the previous parts and summarizes the process, methodology and outcomes of the research presented in this thesis. Based on the model of *Exploratory Browsing*, we present a catalogue of general characteristics for such browsing interfaces, based on which we analyze the effectiveness of our prototypes. We also propose several general suggestions for designing interfaces for media collections.

11.1 Summary of the Process

This thesis describes our research of interface design for media collections, with the main outcome of a new model of *Exploratory Browsing* to inform the design of interfaces to support the exploratory experience with media collections.

The process of this research is illustrated in the overview of the thesis structure (see Figure 1.6 in Chapter 1). Inspired by the existing models of searching and browsing, we proposed our **definition and initial model of *Exploratory Browsing*** (cf. Section 1.2 in Chapter 1). According to the users' browsing objectives, we grouped the *Exploratory Browsing* activities into three categories: *Search Browsing*, *General Purpose Browsing* and *Serendipitous Browsing*. *Search Browsing* is a directed and structured activity with a specific target. *General Purpose Browsing* involves an exploration in areas of interest. *Serendipitous Browsing* is unstructured and random. In this thesis, *Exploratory Browsing* refers to the latter two browsing activities which involve no specific goal. In our initial model of *Exploratory Browsing*, these three activities were conceived as interconnected (see Figure 1.3 in Chapter 1).

We investigated the existing work on interface design for media collections (cf. Chapter 2). Our **analysis of some representative applications** showed a lack of interfaces to efficiently support *Exploratory Browsing* in media collections. This motivated us to systematically explore the full range of browsing activities, specifically their underlying transitional relations, aiming at informing a new model for the design of interfaces to enhance the Browsing experience with media collections.

Our initial exploration started with similarity (cf. Chapter 3). We developed *PhotoSim* and *MusicSim* to support similarity-based browsing in **personal media collections**. The preliminary evaluation confirmed an overall positive feedback towards similarity. However, its performance does not meet the high-level expectations, and users prefer reliable criteria. Thereby, we suggested that similarity should be used as a secondary criterion and offered on demand. In Chapter 4, we explored an effective combination of similarity, commonly used reliable criteria and additional information, specifically User-Generated-Data (UGD). We developed two prototypes *CloudMonster* and *PhotoMagnets* to enhance the browsing flexibility in personal media collections. Beyond the simple manual override provided in *PhotoSim* and *MusicSim*, we proposed more elaborate user interactions. During the evaluation, the afforded browsing flexibility was especially appreciated. The employed magnet metaphor has proven to be an intuitive and efficient assistance for formulating complex query. It is also helpful to conduct more explorative tasks beyond an explicit search, such as grouping, comparing, storytelling and album creation. The results also confirmed the effectiveness of UGD to enhance the browsing flexibility with personal media collections. Through an observation of the users' behavior, we discovered the transitional relations between prominent *General Purpose Browsing* and the other two browsing activities. We then refined the initial model of *Exploratory Browsing* (cf. Section 5.3 in Chapter 5).

Because of the users' overall familiarity with their own collections, *Serendipitous Browsing* was observed not to be prevalent in personal media collections. To investigate *Serendipitous Browsing* activities in detail, we shifted our focus from personal media collections to **online communities**. As the previous chapters have confirmed the effectiveness of UGD to improve the browsing flexibility, we explored the usage of abundant collaborative UGD contributed by large numbers of users in online communities. We first conducted a survey of user behavior and relevant collaborative UGD in those platforms (cf. Chapter 6). The results revealed that the *Exploratory Browsing* behavior exists in such an environment, which can be enhanced by collaborative UGD. By integrating collaborative UGD in the browsing interfaces, we developed four prototypes (cf. Chapter 7, 8, and 9): *HisFlocks* for browsing and comparison of multiple users' tastes, *TagClusters* for improving semantic understanding of collaborative tags, *SARA* for generating personalized package recommendations, and *MusicTrends* for exploring worldwide consumption trends. Building upon the insights from the development and evaluation of these prototypes, we confirmed the effectiveness of collaborative UGD to enhance the *Exploratory Browsing* experience in online communities. We also uncovered the transitional relations between *Serendipitous Browsing* and the other two browsing activities. Based on the discovered full range of interconnections between the browsing activities, we further enriched the model of *Exploratory Browsing*.

11.2 Summary of the Methodology

The objective of this research is to support the *Exploratory Browsing* experience with media collections, which requires extensive user involvement. We applied strategies of **User-Centered design** in our work (cf. Section 1.4 in Chapter 1). During the prototype development, end users were closely involved into the three-phase design process. For the **analysis** of the users' requirements and tasks, we first discussed with experts, which helped to define the target user group and to gain the first impression of their behavior and needs. After carefully selecting the end user group, we recruited target users and conducted a preliminary study with them. A combination of questionnaire, interview and user observation was used to gain insights on the users' characteristics, which helped to achieve a better understanding of the usage scenarios and detailed user requirements. Data was collected through questionnaire and taking video, audio, notes and photographs. To validate the initial findings in this study, we then conducted large-scale survey with broader audience from the target group. Based on the insights derived from the preliminarily test and survey, several initial **design** concepts were developed. They were refined in an iterative process, in which samples of end users were involved. These concepts were mainly evaluated with low-fidelity prototypes, since users may feel free to react and propose alternatives (Sharp et al., 2007). The employed prototypes included sketches, storyboards and paper prototypes (see examples in Figure 1.5 in Chapter 1). During the implementation of the final concepts, quick pilots were employed to help making decisions on specific design issues.

After fully implementing the prototype, we evaluated it in a **user study**, which was video recorded using the Think-Aloud protocol. For all questionnaires and interviews, we took pilot test in advance to measure time per participant and to ensure a smooth workflow, while avoiding possible ambiguity and misunderstanding. The study consisted of a pre-questionnaire, an interview and a post-questionnaire. At the beginning of the study, the participants filled out a pre-questionnaire about their demographic information and general experience with the issue explored in the prototype. Then they were given a brief introduction of the prototype and encouraged to play around with it. After this trial session, they joined an interview. The user study was concluded with a post-questionnaire concerning the overall impression of the tested system(s).

Considering the different characteristics of the prototypes, we applied different interview approaches and accordingly different methods to analyze the collected data. As Figure 11.1 shows, we used 4 different categories of evaluation methods. **User discussion** helps to judge the suitability of initial concepts in their early stage. For example, the users' feedback towards the similarity-based interfaces in *PhotoSim* and *MusicSim* (cf. Section 3.1.2 and 3.2.2 in Chapter 3 respectively) motivated us to combine similarity with reliable metadata and additional information, which led to the other two prototypes *CloudMonster* and *PhotoMagnets* (cf. Chapter 4).

System Characteristic	Approach		Prototype	Collected data				Remarks on the analysis
				QD	QF	UR	UO	
Initial concept	User discussion		PhotoSim MusicSim		√		√	
Novel concept	Empirical study		Cloud Monster		√	√	√	
	Empirical and explorative studies		Photo Magnets		√	√	√	-QF: interviewed both hobby and professional users
Comparability	Comparative study	Between-system	HisFlocks	√	√		√	
			Tag Clusters	√	√	√	√	-QD: time and precision -UR: task easiness and system helpfulness
		Within-system	SARA	√	√	√	√	-QD: time and plan quality -UR: perceived time, plan quality, mode preference -User grouping
Free exploration	Insight-based evaluation		Music Trends	√	√	√	√	-QD: time, number of insights, insight category and quality (spontaneity and complexity) -User grouping

QD: quantitative data, QF: qualitative feedback, UR: user rating, UO: user observation

Figure 11.1: Different evaluation approaches used in our research.

For a rather novel interface, instead of comparing with another quite different system, we chose an in-depth analysis, which addresses specifically the users' quantitative feedback and their behavior. We conducted an **empirical study** with *CloudMonster* (cf. Section 4.1.3 in Chapter 4) based on a set of pre-defined tasks, which were selected based on several considerations, such as task simplicity, coverage of main functionalities and reasonable completion time. Although this method helps to evaluate the system usability, especially with the key interface components, such as multiple initial views and magnets in *CloudMonster*, it may constrain the users' practical usages. Therefore, in the user study of *PhotoMagnets*, we combined empirical evaluation with **explorative study** (cf. Section 4.2.3 in Chapter 4). To encourage the users to explore the system more freely and to enable a closer observation of their behavior, we included two open-ended tasks of storytelling and album creation. The participants were better motivated to explore the tool extensively and discovered several new usages especially with magnets, which were not expect in our initial design, such as comprehending the implied social information and the discovery of negative magnets. Besides studying hobby photographers, we further discussed *PhotoMagnets* with experts, and their positive feedback indicated a potential usage of this tool in a professional user group.

For an improvement of existing systems, we chose a **between-system comparison**, which helped to analyze the advantages and limitations of different systems. For example, we compared *HisFlocks* with Last.fm musical compatibility bar and Last.fm Explorer (cf. Section 7.1 in Chapter 7), and *HisFlocks* was perceived as a more efficient assistance for browsing and comparing users' musical tastes in detail. We also applied this comparative strategy in the evaluation of *TagClusters*, a semantic aggregation of TagClouds (cf. Section 8.1.3 in Chapter

8). A set of pre-defined tasks were carefully selected, covering the main features of both systems in the aspects of search, browsing, comparison and relation understanding. The results revealed that *TagClusters* have an advantage in supporting a topic-oriented exploration.

Besides comparing different systems, this comparative approach can be applied to evaluate different **variations of the same system**. For example, we evaluated *SARA* with its three recommendation modes, and the users were required to create a plan with each of these modes (cf. Section 8.2.5 in Chapter 8). Combining both quantitative and qualitative results in the analysis eases the understanding of specific issues. For example, the conflict between objectively measured time and the users' perceived time for the local recommendation mode indicated that the system effectiveness is only partially determined by efficiency, and that active user control can efficiently improve user satisfaction. We measured the quality of the final plans and discovered that the quantitative criteria could not efficiently reflect how these plans satisfied the users' dynamically changing requirements. Therefore, we relied more on the users' qualitative feedback. Two rounds of user ratings for the plan quality revealed a subjective bias towards those plans, in which the users invested more effort. As the users' scores tend to become more objective with the decrease of novelty, we suggested employing an anonymous evaluation for personal solutions. According to the users' characteristics, we grouped them into different groups and revealed a possible correlation between mode preference and the user's genre and trip planning experience.

For an interface involving no specific tasks or goals, we applied an **insight-based evaluation**. In the user study of *MusicTrends* (cf. Section 9.1.3 in Chapter 9), the users were encouraged to discover insights in a free exploration, without requirement for speed or quantity. As the objective insights can be quantitatively measured, we proposed two metrics to analyze the insights derived by the users. An insight distribution was achieved through an insight classification, which helped to gain an overall impression of the system's effectiveness to support deriving different category of insights. We measured the quality of insight by its spontaneity and complexity. We then analyzed user performance based on number of insights, time per insight and insight quality. Similar to *SARA*, we applied a user grouping and derived a possible correlation between the users' performance and their Last.fm experience.

We especially highlighted the value of **open-ended tasks** and **free exploration**, which motivated the user to use the system more extensively and enabled us to achieve a close observation of their behavior. Regarding the analysis of the collected data, we focused more on the qualitative data in the user discussion and in-depth study. For the comparative studies, we combined both quantitative and qualitative analysis. **Grouping users into different groups** according to their characteristics, such as gender and experience, can facilitate investigating possible correlation between the users' performance and their characteristics. We used different criteria to measure the **quality of the output**. For a personal solution such as a trip plan in *SARA*, we relied more on the users' qualitative feedback. For objective output, such as insights gained in an exploration in *MusicTrends*, we analyzed the quality based on quantitative metrics, such as spontaneity and complexity. Overall, combining quantitative data with users' ratings and qualitative feedback can offer a thorough impression of the system effectiveness. Users' comments and suggestions also help to gather insights on system improvement. We observed user behavior closely in the evaluation of all prototypes, which allowed us obtaining a better understanding of the user's practical activities, and thus informed reflections on the model of *Exploratory Browsing*.

11.3 Summary of the Results

One of the main contributions of this thesis is a new model of *Exploratory Browsing* to inform the interface design to enhance the browsing experience with media collections. We proposed the initial model of *Exploratory Browsing* in Chapter 1 (cf. Figure 1.3 in Chapter 1). Through the exploration with personal media collections in Part II, we achieved a better understanding especially with the transitional relations between *General Purpose Browsing* and the other two browsing activities. Accordingly, we refined the initial model of *Exploratory Browsing* (see Figure 5.1 in Chapter 5). In Part III we investigated this model in the context of online communities. We discovered the transitional relations between *Serendipitous Browsing* and the other two browsing activities, which introduced a further enrichment of this model. Through uncovering the full range of transitional relations, we achieved a model of Exploratory Browsing, in which the three main browsing activities are interconnected (see Figure 11.2).

Based on this model we present a catalogue of general interface characteristics to support *Exploratory Browsing*. We utilize this catalogue as criteria to analyze the effectiveness of our prototypes. We also propose several general suggestions for designing interfaces for media collections.

11.3.1 The Model of Exploratory Browsing

We defined *Exploratory Browsing* as the behavior when the user has no specific targets and needs to discover areas of interest (exploratory), in which she or he can explore in detail and possibly discover acceptable items (browsing) (cf. Section 1.2). We categorized browsing activities as *Search Browsing*, *General Purpose Browsing* and *Serendipitous Browsing*, which are interconnected as Figure 11.2 illustrates. *Exploratory Browsing* refers to the latter two browsing activities, which go beyond the explicit search with specific targets.

The users' browsing strategies are largely determined by their intensions. With well-defined targets, users often start with *Search Browsing* (A). In online communities where large numbers of items are available, users tend to locate the familiar ones as the starting points. They conduct explicit search, navigate through the collection, or jump directly to the target. If users have a loosely defined goal or have not enough knowledge to formulate their requirements, they are likely to enter the phase of *General Purpose Browsing* (B), in which they explore relevant neighborhood in detail and possibly find some satisfying items. As discussed in Chapter 5 and 10, *General Purpose Browsing* is prevalent in both personal collections and online communities, which is consistent with the discussion of user behavior in Chapter 2: People seldom look for one specific item, and instead a significantly larger proportion of time is spent in browsing activities, where they can find the matches on the way of exploration. In an area of interest, users may invest effort in comparing, comprehending and interpreting the underlying relations between items. Without any objective, they may engage in a free exploration, in which they heavily rely on serendipity (H). *Serendipitous Browsing* is especially popular in online communities, because of the users' overall lower confidence with large collections of items and their abundant accompanying metadata. Serendipity can stimulate curiosity and encourage exploration. It facilitates discovering new items, encourage users to revisit older collections and rediscover forgotten items.

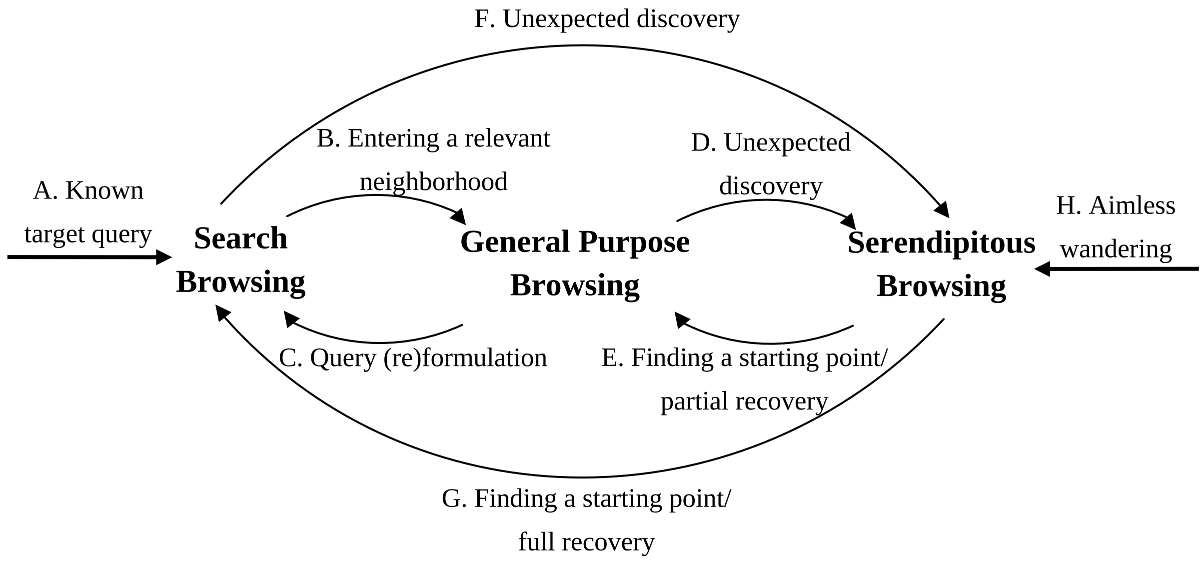


Figure 11.2: The model of Exploratory Browsing.

As Figure 11.2 shows, these three main browsing activities are interconnected. On the way of browsing, users' requirements may become clear and thus leading to a query (re)formulation (C). On the other hand, their attention may be distracted by unexpected but interesting findings, thus enter the phase of *Serendipitous Browsing* (D, F). This is especially the case in online communities, where the users' focus may dynamically change when facing many alternatives. During the free exploration, users may discover an interesting area or recall a certain item, and then start to explore the relevant neighborhood or explicitly look for this specific item. Accordingly, the browsing activity will transfer to *General Purpose Browsing* (E) or *Search Browsing* (G). In the situations when nothing satisfying is found, or users feel bored or being diverted too far way from the original intentions, they may recover from serendipity and return back to a status reached previously (E, G).

11.3.2 General Interface Characteristics to Support Exploratory Browsing

Based on the initial model of *Exploratory Browsing* (see Figure 1.3 in Chapter 1), we proposed an initial catalogue of general interfaces characteristics and utilized it to analyze some representative applications (see Figure 2.11 and 2.12 in Chapter 2).

Through the exploration in Part II and III, we achieved a better understanding of the users' browsing behavior and proposed a refined model of *Exploratory Browsing* (see Figure 11.2). Based on this model, especially the uncovered transitional relations between the browsing activities, we enrich the initial catalogue of general interfaces characteristics (see Figure 11.3). We employ this catalogue as criteria to analyze the effectiveness of our prototypes.

Characteristic \ Prototype	Photo Sim	Music Sim	Cloud Monster	Photo Magnets	His Flocks	Tag Clusters	SARA	Music Trends
Search Browsing								
Structured navigation	√			√	√	√	√	√
Direct jump	√			√	√		√	√
Keyword-based search		√	√	√		√		
Tagging facility				√				
Exploratory Browsing								
General Purpose Browsing								
Overview	√	√	√	√	√			
Defining areas of interest				√	√		√	
Multiple representations			√	√			√	√
Multi-dimensional exploration			√	√	√		√	√
Serendipitous Browsing								
Nonlinear representation	√	√	√	√	√	√	√	√
Steerable representation	√	√	√	√	√		√	
Backtracking	√	√	√	√	√		√	√
Introducing unfamiliarity	√	√	√	√	√	√	√	√

Figure 11.3: The catalogue of general interface characteristics to support Exploratory Browsing.

Search Browsing

The objective of *Search Browsing* is to find specific targets. Therefore, **keyword-based search** should be allowed, which facilitates quickly locating these targets, especially in a rather large graph. For items with meaningful names, such as music, an explicit query can be efficiently conducted with standard keywords, such as artist, album and song. A **tagging**

functionality can be a compensation for collections which initially lack meaningful names, such as photos. As tagging is often considered tedious and time-consuming, a convenient and enjoyable tagging facility, such as the one offered in *PhotoMagnets*, is essential to encourage user contribution. Since users may have difficulties to formulate their requirements, an efficient facility for query (re)formulation is necessary. The magnet metaphor was proved in Part II an efficient assistance, which allows the user actively defining and refining a complex query with elaborate interactions. A prediction of the query results can help the user making timely adjustment. For example, in PhotoMangets, when a magnet is swiped over the timeline, events are rendered differently according to the number of matches in each of them, which facilitate the user to make final selections.

Besides explicit query, users may navigate through a collection to locate a known target. Therefore, a **structured navigation** should be allowed. According to the characteristics of the collections, their structure can be represented differently. A hierarchical structure can facilitate gaining an overall impression of the entire collection and also help to locate a certain sub-collection. For photo collections, this hierarchy, such as the tree view offered in *PhotoSim* and the event structure provided in *PhotoMagnets*, can be extracted from the camera's time stamps or the folder structure. In *TagClusters*, the hierarchical structure is extracted from a semantic analysis of tags. *SARA* provides several sight categories, in which sights are ordered by their popularity. Similar to photos' captured time, the users' consumption records bring another kind of consumption time stamps, which allow displaying consumed items as a sequence, such as the ones offered in *HisFlocks* and *MusicTrends*. These structures can also be used to quickly locate a specific target. The user can **jump directly** to a specific area, such as a specific event or certain time period.

General Purpose Browsing

For a possibly large collection, **overview** is important to help gaining an overall impression quickly. *PhotoSim*, *MusicSim* and *CloudMonster* offer an overview for the entire collection. For a time-centric system, such as *HisFlocks*, an overview on the entire timeline can be achieved by setting the time interval as all time frames. A hierarchical structure, such as the tree view in *PhotoSim* and the timeline in *PhotoMagnets*, can also be used as a summary for the entire collection.

To facilitate *General Purpose Browsing*, it is essential to allow **defining areas of interest**. The user can define a time span in *HisFlocks* by adjusting the interval for the time slider. *PhotoMagnets* allows defining areas of interest by loading multiple events. In *SARA* the user can filter preferred sight category with the sliders for sight preference.

Multiple representations can offer different views of the same collection and thus assist exploration along different criteria. This concept has been widely applied in the field of Information Visualization (InfoVis) and also introduced in our prototypes. *MusicTrends* and *SARA* display information both in a list and on a map. *CloudMonster* provides three views for the entire collection: In genre view, all songs are grouped into genre sectors. Similarity view places songs on a map according to their similarity. In popularity view, songs are represented as a spiral and the most popular ones are placed in the center. *PhotoMagnets* displays events hierarchically in a tree view or chronologically in bubbles. To facilitate a seamless switch between these views, these representations should be coordinated closely (which is also called multiple coordinated views in InfoVis), for example connected by the focus item.

Individual representations offer linear scan along respective criterion. As the users' metal model of the collections often goes beyond the simply linear scan, **multi-dimensional exploration** should be allowed to facilitate browsing and comparison along multiple principles. Magnets enable a query based on multiple criteria and the dynamically adjustable representation allows a flexible exploration in the retrieval results. Both spatial and temporal dimension are integrated in *HisFlocks*, *SARA* and *MusicTrends*, which allows exploring temporal changes in certain area.

Serendipitous Browsing

A **nonlinear representation** can encourage discovering new relations between items. Map-based view is extensively used in our prototypes. Items can be displayed on such a view based on their content similarity (in *PhotoSim* and *MusicSim*), semantic relatedness (in *HisFlocks* and *TagClusters*), or geographic locations (in *SARA* and *MusicTrends*). The magnet metaphor is another option to produces a flexible representation for the query results.

To stimulate unexpected but interesting findings, users should be allowed to **steer the representation** through interactions. The key parameters for producing the representation should be adjustable, such as the similarity threshold and number of clusters in *PhotoSim* and *MusicSim*, magnet magnitude in *CloudMonster* and *PhotoMagnets*, font size and color in *TagClusters*, time interval in *HisFlocks*, and the sliders for sight and route preferences in *SARA*. Beyond adjusting these parameters, elaborate user interactions, such as the magnet metaphor, can further encourage an active exploration in a dynamically adjustable interface.

Backtracking is a common strategy in interface design to cancel undesired actions. In the case of media collections, it is also essential to facilitate recovering from serendipity. A simple solution is an "undo" functionality for interaction mistakes, such as the ones provided in *PhotoSim*, *MusicSim* and *CloudMonster*. Resetting the parameters, such as the preference sliders in *SARA* and the time slider in *HisFlocks* and *MusicTrends*, make it easy to switch back to a previous status by reproducing an exact same representation. The users' usage history can also be used as an anchor point back onto the original track, by saving the current status for future references, as provided in *CloudMonster* and *PhotoMagnets*.

Besides the functionalities discussed above, serendipity can be induced by **unfamiliarity**. One simple method to stimulate serendipity is to enlarge the exploration neighborhood, for example, by presenting less relevant items. Although the performance of content analysis is not always as precise as expected, it stimulates serendipity indeed by lower similarity. User-Generated Data (UGD) can also encourage discovering new findings, which users may not be aware of, and thus create another kind of serendipity. For example, user-generated tags can reflect relevant social relations between photos. Collaborative UGD can further encourage serendipity by introducing large sets of data created by others, which encourage new discoveries in a broader and more diverse range.

Based on the general characteristics proposed above, we analyze the effectiveness of our prototypes. When looking at Figure 11.3, one can observe that *PhotoMagnets* outperforms other prototypes in supporting all the suggested characteristics. *Exploratory Browsing*, especially *General Purpose Browsing*, is poorly supported in *TagClusters*. The effectiveness of other prototypes in supporting *Search Browsing* is comparative. Concerning *General Purpose Browsing*, *PhotoMagnets*, *CloudMonster*, *SARA* and *HisFlocks* meet the requirements better. All prototypes support *Serendipitous Browsing* generally well.

11.3.3 General Suggestions for Interface Design

Based on our experience accumulated in designing interfaces to support *Exploratory Browsing*, we have identified several general issues which should be considered in designing interfaces for media collection. Specifically we categorize them into five categories (see Figure 11.4).

Prototype / **Characteristic**	**Photo Sim**	**Music Sim**	**Cloud Monster**	**Photo Magnets**	**His Flocks**	**Tag Clusters**	**SARA**	**Music Trends**
Using different data sources								
Original metadata	√	√	√	√	√	√	√	√
System-generated data	√	√	√	√			√	
User-generated data			√	√	√	√	√	√
Flexibility								
Multiple representations			√	√	√		√	√
User interaction	√	√	√	√	√	√	√	√
Personalization	√	√	√	√	√	√	√	
Understandability	√	√	√	√	√	√	√	√
Reproducibility	√	√	√	√	√	√	√	√
Scalability	√	√			√	√		

Figure 11.4: General suggestions for designing interfaces for media collections.

Using Different Data Sources

Data with different forms and origins can be combined to enrich the browsing possibilities. Normally the most reliable information comes from the item's **original metadata**, such as captured time of photos and music's ID3tags of artist and album. Users often use these fundamental and reliable criteria as the main principles for organization and browsing, therefore we introduce them in all prototypes. Besides these commonly used reliable criteria, additional options are also appreciated, which can be achieved by utilizing the system's and users' contributions. Additional information can be obtained through an analysis of the raw data, such as events extracted from photos' time stamps in *PhotoMagnets*, and hierarchical tag structure built on a text analysis in *TagClusters*. Using content analysis, additional features besides similarity can be read out, such as background color and light status in photos, and gender, genre and mood for music. They can be used to create additional views for the entire collection, to conduct example-based search, or to generate content-based suggestions. Although this kind of **system-generated data** is less reliable, it complements the browsing possibilities offered by commonly used reliable criteria. **User-Generated Data (UGD)** is another source of additional information, which is extensively used in our prototypes. Data generated by single users conveys semantic relatedness of items on a personal level. When large numbers of users are involved, this information becomes more general and reliable in a broader range. We categorize this collaborative UGD as implicit and explicit UGD. Implicit UGD includes data recorded implicitly through user interactions, such as consumption histories, which reflects the users' personal tastes and consumption patterns. Explicit UGD is contributed by users on purpose, such as tags, ratings and reviews, which conveys the users' opinions and experience in detail. Since the size of a collaborative UGD collection is normally quite large, an aggregation is necessary to reduce the data complexity. Our exploration in Part II and III confirmed the effectiveness of UGD to encourage discovering new items, deriving additional insights and acquire knowledge which cannot be easily obtained through a linear scan of respective collections.

Flexibility

People's activities around media collections are highly dynamic and often unstructured (cf. Chapter 2). To support the behavioral dynamics, the interface should allow a high degree of flexibility, which we believe can be introduced in three ways: As users' mental model of their collections often goes beyond a fixed linear structure, **multiple representations** allow them to look at the collections in different ways, thus enhance the browsing flexibility (more details see "*General Purpose Browsing*" in Section 11.3.2). Moreover, different criteria can be combined in one representation to offer more browsing possibilities. An efficient combination with traditional interface elements can accelerate the users' acquaintance with novel concepts. For example, we combined the magnet metaphor with the conventional tree view in *PhotoMagnets*, and the results of the user study confirmed a high degree of interface appropriation. Users may frequently switch back and forth between different views, therefore, they should be closely coordinated and smooth transitions between them should be ensured.

User interaction can help users to make more sense of the underlying system logic. Options to override the results of automation can enhance the degree of **personalization**, and thus improve the users' feeling of control and enjoyment. Especially in a large-scale collection, personalization is essential to support an efficient browsing based on personal preferences and interests. Better motivated users may spontaneously discovery new features of the system which are not foreseen in the original design, such as the advanced usages of magnets in *PhotoMagnets*. For systems built based on content analysis, there is always a semantic gap between low-level features and high-level perception. In this case, users' adjustment offers an opportunity to correct the system's inappropriate interpretations. The simplest way of user manipulation is to allow making changes manually, for example, repositioning items by drag-and-drop. As the size of current collections is always in steep growth, manual adjustment may be tedious and time-consuming. Therefore, it seems more appropriate to offer semi-automatic solutions which efficiently combine the system's intelligence and the users' feedback: Users select several key points and the system conduct the corresponding adjustment. For example, in *PhotoSim* and *MusicSim*, the user can adjust the number of clusters, and accordingly the system will re-cluster items to produce a new layout. Similarly, in the global recommendation mode in *SARA*, the user can replace any sight in a plan and the system will automatically update the subsequent part of this plan. Adjustable representation not only facilitates personalization, but also creates a chance for backtracking. For example, it is possible to recovery from an operation mistake or switch back to a previous status by resetting the corresponding parameters. To improve the directedness and continuity of the traditional turn-taking interactions and offer users a seamless exploratory experience, a synchronous feedback should be tightly coupled with the users' manipulation. Dynamic Query (Ahlberg and Shneiderman, 1994a) is a simple and efficient solution, with which users can observe an immediate change with their actions, for example dragging a magnet or a slider.

Understandability

Understandability is a general requirement for user interfaces. To illustrate the system logic, for example how a representation is organized or why certain items are recommended, additional information can be integrated in the **item rendering** to enhance the understandability. In general, size and color are the most commonly used attributes. Size can represent the popularity of an item, and color can be used to distinguish different categories, or to describe the similarity between items.

The location information can reflect the semantic or geographic relations between items, thus make the **item positioning** understandable and meaningful. Specific concern should be paid to a placement based on overall similarity, which is in some cases arbitrary and unintuitive. We suggest applying similarity in local levels to achieve a more accurate performance, or combine it with other more reliable criteria to offer a clearer explanation of the organization.

Reproducibility

Good reproducibility is essential to maintain the representation consistency. Map-based systems bear good reproducibility with precise geographic location information. Force-directed layout is widely used to automatically group similar items based on their overall similarity. However, this placement is rather arbitrary, due to the randomness of the underlying force simulation mechanism. When the application re-launches, an exact same visualization cannot be promising. This problem can be solved by automatically saving the position of each item. A more precise positioning method beyond the overall similarity should be included to enhance the system reproducibility.

The reproducibility in one usage session can be improved by automatically or manually saving the current status, which enables switching back to a previous status. A same effect can also be achieved by adjusting relevant parameters.

Scalability

The system's scalability is important towards realistic collection sizes. Although map-based interfaces bear good reproducibility with the precise location values, scalability is a crucial issue, especially when many items share similar locations. For systems built based on overall similarity, an overlapping detection mechanism can efficiently solve the problem of occlusion by placing overlapped items further away from each other. It should be combined with a tightly coupling method to avoid possibly large wasted space. Grouping similar items and expanding them on demand can be a general compensation for scalability.

Based on the five criteria discussed above, we analyzed our prototypes accordingly (see Figure 11.4). All prototypes used at least two categories of data sources. More than half of them provide multiple representations. All of them preserve users active control, and allows them actively steering the system and thus creating personalized representations. Items in each system are encoded with additional information. Reproducibility in one usage session is ensured by saving the current status. Due to the randomness of force-directed layout, the exact same layout for two subsequent applications cannot be promising, and our current solution is to save the item locations automatically. Admittedly, a more precise positioning method is necessary. Occlusion problem is solved by an overlapping detection method, but the rather loose representation needs further optimization, and a tight coupling mechanism can make better usage of the space. Considering the limited space between items, such as sights on the *SARA* map, or magnets in *CloudMonster* and *PhotoMagnets*, a semantic grouping can alleviate the problem of occlusion.

12

Future Work

Building upon our experience accumulated in this research, we identify several aspects of interest for improvement and future work. The prototypes presented in this thesis can be improved with additional refinement. In this chapter, we focus more on the aspects of interest for our future research as well as issues worth consideration for the area of interface design to support exploratory experience.

12.1 The Model of Exploratory Browsing

Our exploration in this thesis helps deepening the understanding of how people's behavior and requirements have an influence on the design decisions. We have proposed a new model of *Exploratory Browsing* to inform the interface design to support the *Exploratory Browsing* experience with media collections. Our studies yield promising reflections on the refinement of the model, which brings substantial insights for designing browsing interfaces for media collections. This model needs more extensive investigation, especially with the stimulators to ensure a smooth transition between different browsing activities.

One direction in our future work is to elaborate the current model of *Exploratory Browsing* to guide the design of new interfaces for media collections. In this thesis we have explored two main usage contexts: browsing with personal collections and in online communities. To improve the generality of our model, co-located environment needs investigation, which differs from the other two usage contexts in the social aspect. It would be interesting to observe people's behavior in a co-presented environment, and to examine how their exploratory activities would differ from those in the other two usage cases. We expect that such an exploration can bring substantial insights to generalize the model of *Exploratory Browsing* by covering broader usage contexts.

Another aspect worth consideration is the evaluation methodology. Evaluating exploratory interfaces is difficult, and the main challenge is to find ways to evaluate the system's effectiveness to support the user's exploratory experience. To understand the user behavior and how it has impact on the interface design, we have applied different evaluation approaches, according to the different characteristics of our prototypes (cf. Figure 11.1 in Chapter 11), for example, empirical evaluation and explorative study, comparative study and insight-based evaluation. The experimental studies reported in this thesis were conducted in a laboratory environment. Long-term studies can provide more in-depth comprehension of system usability (Shneiderman and Plaisant, 2006). Also, deploying real tasks rather than synthetic ones can encourage the users' engagement (Saraiya et al., 2005). We expect that such exploratory and long-term studies could be more informative and bring more values to examine whether the users' practical browsing activities can be efficiently supported and how they inform possible refinement of the current model of *Exploratory Browsing*.

Through such investigation and experimentation, we expect to propose a more elaborate model of *Exploratory Browsing*. All in all, we hope that our work can shed some light on the interface design by informing the model of *Exploratory Browsing* and corresponding general guidelines to support the exploratory experience with media collections.

12.2 Inspirations for the Relevant Domains

Besides designing interfaces for media collections, we would like to explore the applicability of the model of *Exploratory Browsing* in relevant domains. *Exploratory Browsing* behavior exists prevalently in people's daily life, varying from looking for a nice picture in a personal collection to buying a good travel guidebook online. Because of the overall familiarity with personal collections, users play active role in deciding which area to explore and how the relevant items should be organized. In online communities where abundant information is available, users tend to rely more on the suggestions from the system or other users. In such an environment, the system acts more like a Recommender System, which offers appropriate products and services according to the users' requirements.

To do so, users' exploratory behavior, their needs and tasks need systematic exploration. In the domain of e-commerce, researchers have explored people's online shopping behavior. Several studies focused on the analysis of the customers' information seeking patterns, which highlighted the necessity to support such exploratory activities (Detlor et al., 2003; Bauernfeind and Zins, 2006). Based on the observation of customers' navigational pattern in online shops, Moe (2007) categorized online shopper's objective into three types: searching, browsing and knowledge-building. The ability to categorize customers' objectives is emphasized for identifying potential buyers and designing more effective and customized services to stimulate purchase. Although there are several studies on the online shopper's exploratory behavior, systematic exploration on how to reflect these findings on interface design is still in an infant phase.

As one can observe, there is a certain similarity between people's exploratory behavior with media collections and in other contexts, a point of interest for us is to examine our concept of *Exploratory Browsing* in these relevant domains. For example, in the development and study of *SARA*, an Incremental Travel Recommender System (cf. Section 8.2 in Chapter 8), we have explored people's exploratory trip planning behavior in the tourism domain. We believe that such an investigation will inform the possibility and suitability for transferring the model of *Exploratory Browsing* for media collections to other relevant research areas.

Bibliography

Aamodt A. and Plaza E. (1994). Case-based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. Artificial Intelligence Communications, 7(1):39-59.

Ahlberg C. and Shneiderman B. (1994a). Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In Proceedings of the SIGCHI conference on Human factors in computing systems (Boston, Massachusetts, USA, April 24-28, 1994). CHI 1994. ACM Press, New York, NY, USA, 313-317.

Ahlberg C. and Shneiderman B. (1994b). Visual Information Seeking Using the FilmFinder. In Proceedings of the SIGCHI conference on Human factors in computing systems (Boston, Massachusetts, USA, April 24-28, 1994). CHI 1994. ACM Press, New York, NY, USA, 433-434.

Ames M. and Manguy L. (2006). PhotoArcs: A Tool for Creating and Sharing Photo-narratives. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ext. Abstracts (Quebec, Canada, April 22-27, 2006). CHI 2006. ACM Press, New York, NY, USA, 466-471.

Ames M., and Naaman M. (2007). Why We Tag: Motivations for Annotation in Mobile and Online Media. In proceedings of the SIGCHI conference on Human factors in computing systems (San Jose, California, USA, April 28-May 3, 2007). CHI 2007. ACM Press, New York, NY, USA, 971-980.

Andric A. and Haus G. (2006). Automatic Playlist Generation based on Tracking Users Listening Habits. Journal of Multimedia Tools and Application, 29(2): 127-151.

Arjan R., Pfeil U., Zaphiris P. (2008). Age Differences in Online Social Networking. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ext. Abstracts (Florence, Italy, April 5-10, 2008). CHI 2008. ACM Press, New York, NY, USA, 2739-2744.

Aucouturier J. J. and Pachet F. (2002). Music Similarity Measures: What's the Use? In Proceedings of the International Society for Music Information Retrieval (Paris, France, October 13-17, 2002). ISMIR 2002.

Bainbridge D., Cunninghan S. J., Downie J. S. (2003). How do People Describe Their Music Information Needs. In Proceedings of the International Society for Music Information Retrieval (Baltimore, Maryland, USA, October 26-30, 2003). ISMIR 2003.

Balabanović M., Chu L. L., Wolff G. J. (2000). Storytelling with Digital Photographs. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ext. Abstracts (The Hague, The Netherlands, April 1-6, 2000). CHI 2000. ACM Press, New York, NY, USA, 2657-2660.

Bauernfeind U. and Zins A. H. (2006). The Perception of Exploratory Browsing and Trust with Recommender Websites. Information technology & tourism, 8(2): 121-136.

Baumann S., Bassoli A., Jung B., Wisniowski M. (2007). BluetunA: Let Your Neighbor Know What Music You Like. In Proceedings of the SIGCHI conference on Human factors in computing systems, Ext. Abstracts (San Jose, California, USA, April 28-May 3, 2007). CHI 2007. ACM Press, New York, NY, USA, 1941-1946.

Baumgartner H. and Steenkamp J-B E. M. (1996). Exploratory Consumer Buying Behavior: Conceptualization and Measurement. International Journal of Research in Marketing, 13(2):121-137.

Baur D., Langer T., Butz A. 2009. Shades of Music: Letting Users Discover Sub-Song Similarities. In Proceedings of the International Society for Music Information Retrieval (Kobe, Japan, October 26-30, 2009). ISMIR 2009. 111-115.

Bederson B. B. (2001). PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps. In Proceedings of the 14th annual ACM Symposium on User interface Software and Technology (Orlando, FL, USA, November 11-14, 2001). UIST 2001. ACM Press, New York, NY, USA, 71–80.

Bederson B. B., Grosjean J., Meyer J. (2004). Toolkit design for interactive structured graphics. IEEE Transactions on software engineering, 30:535–546.

Begelman G., Keller P., Smadja F. (2006). Automated Tag Clustering: Improving Search and Exploration in the Tag Space. In Proceedings of International World Wide Web Conference (Edinburgh, UK, May 22-26, 2006). WWW 2006. ACM Press, New York, NY, USA.

Bentley F., Metcalf C., Harboe G. (2006). Personal vs. Commercial Content: The Similarities between Consumer Use of Photos and Music. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Quebec, Canada, April 22-27, 2006). CHI 2006. ACM Press, New York, NY, USA, 667-676.

Bilgic M. and Mooney R. J. (2005). Explaining Recommendations: Satisfaction vs. Promotion. In Proceedings of the International Conference on Intelligent User Interfaces (San Diego, California, USA, January 10-13, 2005). IUI 2005. ACM Press, New York, NY, USA, 13-18.

Boll S., Sandhaus P., Scherp A., Thieme S. (2007). MetaXa-Context-and Content-driven Metadata Enhancement for Personal Photo Books. In Proceedings of the 13th International Multimedia Modeling Conference (Singapore, January 9-12, 2007). MMM 2007. Springer Verlag, Berlin/Heidelberg, Germany, 332-343.

Brown B. A. T., Geelhoed E., Sellen A. J. (2001a). The Use of Conventional and New Music Media: Implications for Future Technologies. HP Lab technical report, HPL-2001-102.

Brown B. A. T., Sellen A. J., Geelhoed E. (2001b). Music Sharing as Computer Supported Collaborative Application. In Proceedings of the seventh conference on European Conference on Computer Supported Cooperative Work (Bonn, Germany, September 16-20, 2001). ECSCW 2001. ACM Press, New York, NY, USA, 179-198.

Callery A. (1996). Yahoo! Cataloguing the Web. http://www.library.ucsb.edu/untangle/callery.html - cited 2010/5/27.

Campos J. and de Figueiredo A. D. (2001). Searching the Unsearchable: Inducing Serendipitous Insights. In Proceedings of the workshop Program at the Fourth International Conference on Case-based Reasoning (Vancouver, BC, Canada, July 30-August 2, 2001). ICCBR 2001. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.20.7895&rep=rep1&type=pdf- cited 2010/5/27.

Canny J. (2002). Collaborative Filtering with Privacy via Factor Analysis. In Proceedings of the international ACM SIGIR conference on Research and development in information retrieval (Tampere, Finland, August 11-15, 2002). SIGIR 2002. ACM Press, New York, NY, USA, 238-245.

Card S. K., Mackinlay J. D., Shneiderman B. (1999). Readings in Information Visualization, Using Visualization to Think. Morgan Kaufmann, San Fransisco, CA, USA.

Carson C., Thomas M., Belongie S., Hellerstein J. M., Malik J. (1999). Blobworld: A System for Region-based Image Indexing and Retrieval. In Proceedings of the 3rd International Conference on Visual Information and Information Systems (Amsterdam, The Netherlands, June 2-4, 1999). VISUAL 1999. Springer Verlag, Berlin/Heidelberg, Germany, 660.

Cha M., Kwak H., Rodriguez P., Ahn Y. Y., Moon S. (2007). I Tube, You Tube, Everybody Tubes: Analyzing the World’s Largest User Generated Content Video System. In Proceedings of the Internet Measurement Conference (San Diego, CA, USA, October 24-26, 2007). IMC 2007. ACM Press, New York, NY, USA, 1-14.

Chatzichristofis S. A. and Boutalis Y. S. (2008). FCTH: Fuzzy Color and Texture Histogram – A Low Level Feature for Accurate Image Retrival. In Proceedings of the 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services (Klagenfurt, Austria, May 7-9, 2008). WIAMIS 2008. IEEE Computer Society, 191-196.

Chen Y. X., Baur D., Butz A. (2010a). Gaining Musical Insights: Visualizing Multiple Listening Histories. To appear in Proceedings of the international workshop on Visual Interfaces to the Social and Semantic Web (Hong Kong, China, February 7, 2010). VISSW 2010. CEUR-WS.org.

Chen Y. X., Boring S., Butz, A. (2010b). How Last.fm Illustrates the Musical World: User Behavior and Relevant User-Generated Content. To appear in Proceedings of the international workshop on Visual Interfaces to the Social and Semantic Web (Hong Kong, China, February 7, 2010). VISSW 2010. CEUR-WS.org.

Chen Y. X. and Butz A. (2008). *PhotoSim*: Tightly Integrating Image Analysis into a Photo Browsing UI. In Proceedings of 8th International Symposium on Smart Graphics (Rennes, France, August 27-29, 2008). SG 2008. Springer Verlag, Berlin/Heidelberg, Germany, 224-231.

Chen Y. X. and Butz A. (2009). MusicSim: Integrating Audio Analysis and User Feedback in an Interactive Music Browsing UI. In Proceedings of 14th international conference on Intelligent User Interfaces (Sanibel Island, Florida, USA, February 8-11, 2009). IUI 2009. ACM Press, New York, NY, USA, 429-433.

Chen Y. X., Hoyer M., Butz A. (2009a). CloudMonster: Support Flexible Browsing and Searching within Music Collections. In Proceedings of the 12th IFIP TC13 International Conference on Human-Computer Interaction (Uppsala, Sweden, August 26-28, 2009). INTERACT 2009. Springer Verlag, Berlin/Heidelberg, Germany, 428-431.

Chen Y. X., Santamaria R., Butz A., Therón R. (2009b). TagClusters: Semantic Aggregation of Collaborative Tags beyond TagClouds. In Proceedings of 9th International Symposium on Smart Graphics (Salamanca, Spain, May 28-30, 2009). SG 2009. Springer Verlag, Berlin/Heidelberg, Germany, 56-67.

Cheng X., Dale C., Liu J. (2008). Statistics and Social Network of YouTube Videos. In Proceedings of the International Workshop on Quality of Service (Enschede, The Netherlands, June 2-4, 2008). IWQoS 2008. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.150.7896&rep=rep1&type=pdf - cited 2010/5/27.

Cheverst K., Davies N., Mitchell K., Friday A., Efstratiou C. (2000). Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (The Hague, The Netherlands, April 1-6, 2000). CHI 2000. ACM Press, New York, NY, USA, 17-24.

Chi Z., Yan H., Pham T. (1996). Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition. Advances in Fuzzy Systems-Applications and Theory, 10 (1996).

Combs T. T. A., Benderson B. B. (1999). Does Zooming Improve Image Browsing? In Proceedings of the fourth ACM conference on Digital libraries (Berkeley, California, USA, August 11-14, 1999). D-Lib 1999. ACM Press, New York, NY, USA, 130-137.

Cosley D., Lam S. K., Albert I., Konstan J. A., Riedl J. (2003). Is Seeing Believing? How Recommender System Interfaces Affect Users' Opinions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Lauderdale, Florida, USA, April 5-10, 2003). CHI 2003. ACM Press, New York, NY, USA, 585-592.

Cove J. F. and Walsh B. C. (1988). Online Text Retrieval via Browsing. Information Processing & Management, 24(1): 31-37.

Crabtree A., Rodden T., Mariani J. (2004). Collaborating around Collections: Informing the Continued Development of Photoware. In Proceedings of the 2004 ACM conference on Computer Supported Cooperative Work (Chicago, Illinois, November 6-10, 2004). CSCW 2004. ACM Press, New York, NY, USA, 396-405.

Crossen A., Budzik J., Hammond K. J. (2002). Flytrap: Intelligent Group Music Recommendation. In Proceedings of International Conference on Intelligent User

Interfaces (San Francisco, California, USA, January 13-16, 2002). IUI 2002. ACM Press, New York, NY, USA, 184-185.

Cui J., Wen F., Xiao R., Tian Y., Tang X., Tang X. (2007). EasyAlbum: An Interactive Photo Annotation System based on Face Clustering and Re-ranking. In proceedings of the SIGCHI conference on Human factors in computing systems (San Jose, California, USA, April 28-May 3, 2007), CHI 2007. ACM Press, New York, NY, USA, 367-376.

Cui J. and Zhang C. (2007). Combining Stroke-based and Selection-based Relevance Feedback for Content-based Image Retrieval. In Proceedings of the ACM international conference on Multimedia (Augsburg, Germany, September 24-29, 2007). MM 2007. ACM Press, New York, NY, USA, 329-332.

Cunningham S. J., Bainbridge D., Falconer A. (2006). 'More of an Art than a Science': Supporting the Creation of Playlists and Mixes. In Proceedings of the International Society for Music Information Retrieval (Victoria, BC, Canada, October 8-12, 2006). ISMIR 2006.

Cunningham S. J., Reeves N., Britland M. (2003). Ethnographic Study of Music Information Seeking. In Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries (Houston, TX, USA, May 27-31, 2003). JCDL 2003. ACM Press, New York, NY, USA, 5-16.

Cupchik G. C., Rickert M., Mendelson J. (1982). Similarity and Preference Judgment of Musical Stimuli. Scandinavian Journal of Psychology, 23, 273-282.

Davis M., King S., Good N., Sarvas R. (2004). From Context to Content: Leveraging Context to Infer Media Metadata. In Proceedings of the 12th annual ACM international conference on Multimedia (New York, NY, USA, October 10-16, 2004). MM 2004. ACM Press, New York, NY, USA, 188-195.

Davis M., Smith M., Stentiord F., Bamidele A., Canny J., Good N., King S., Janakiraman R. (2006). Using Context and Similarity for Face and Location Identification. In Proceedings of SPIE conference on Internet Imaging (San Jose, CA, USA, April 15-19, 2006). SPIE Digital Library.

De Bono E. (1990). Lateral Thinking. Penguin Books, UK.

Detlor B., Sproule S., Gupta C. (2003). Pre-Purchase Online Information Seeking. Journal of Electronic Commerce Research, 72-84.

Downie J. S. and Cunningham S. J. (2002). Toward a Theory of Music Information Retrieval Queries: System Design Implications. In Proceedings of the International Symposium for Music Information Retrieval (Paris, France, October 13-17, 2002). ISMIR 2002.

Drucker S. M., Wong C., Roseway A., Glenner S. and De Mar S. (2004). Mediabrowser: Reclaiming the Shoebox. In Proceedings of the working conference on Advanced Visual Interfaces (Gallipoli, Italy, May 25-28, 2004). AVI 2004. ACM Press, New York, NY, USA, 433–436.

Duarte F., Benevenuto F., Almeida V., Almeida J. (2007). Geographical Characterization of YouTube: A Latin American View. In Proceedings of the Latin American Web Congress (Santiago de Chille, October-Novermber, 2007). LA-WEB 2007.

Dubinko M., Kumar R., Magnani J., Novak J., Raghavan P., Tomkins A. (2007). Visualizing Tags over Time. ACM Transactions on the Web, 1(7).

English J., Hearst M., Sinha R., Swearingen K., Yee K-P. (2002). Flexible Search and Navigation using Faceted Metadata. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.6.8556&rep=rep1&type=pdf- cited 2010/5/27.

Fesenmaier D. R. and Jeng J. M. (2000). Assessing Structure in the Pleasure Trip Planning Process. Tourism Analysis, 5:13-17.

Firan C. S., Nejdl, W., Paiu R. (2009). The Benefit of Using Tag-based Profiles. In Proceedings of the Latin American Web Congress (Santiago de Chille, October-November, 2007). LA-WEB 2007.

Fodness D. and Murray B. (1999). A Model of Tourist Information Search Behavior. Journal of Travel Research, 37(3): 220-230.

Frohlich D. M. (2004). Audiophotography: Bringing Photos to Life with Sounds. Springer Verlag, New York, NY, USA.

Frohlich D., Kuchinsky A., Pering C., Don A., Ariss S. (2002). Requirement for Photoware. In proceedings of ACM Conference on Computer Supported Cooperative Work (New Orleans, Louisiana, USA, November 16-20, 2002). CSCW 2002. ACM Press, New York, NY, USA, 166-175.

Fujita H. and Arikawa M. (2007). Creating Animation with Personal Photo Collections and Map for Storytelling. In Proceedings of the Euro American conference on Telematics and information systems (Faro, Portugal, May 14-17, 2007). EATIS 2007. ACM Press, New York, NY, USA, No. 1.

Gates C., Subramanian S., Gutwin C. (2006). DJs' Perspectives on Interaction and Awareness in Nightclubs. In Proceedings of the IEEE International workshop on Distributed Intelligent Systems (Prague, Czech Republic, June 15-16, 2006). DIS 2006. 70-79.

Gemmell J., Aris A., Lueder R. (2005). Telling Stories with mylifebits. In Proceedings of IEEE International Conference on Multimedia & Expo (Amsterdam, The Netherlands, July 6-8, 2005). ICME 2005, 1536-1539.

Girgensohn A., Shipman F., Turner T., Wilcox L. (2010). Flexible Access to Photo Libraries via Time, Place, Tags, and Visual Features. In Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries (Gold Coast, Australia, June 21-25, 2010). JCDL 2010. ACM Press, New York, NY.

Goldberg D., Nichols D., Oki B. M., Terry D. (1992). Using Collaborative Filtering to Weave an Information Tapestry. Communications of the ACM, 35(12): 61-70.

Goto M. and Goto T. (2005). Musicream: New Music Playback Interface for Streaming, Sticking, Sorting, and Recalling Musical Pieces. In Proceedings of the International Society for Music Information Retrieval (London, UK, October 11-15, 2005). ISMIR 2005.

Goy A. and Magro D. (2004). Dynamic Configuration of a Personalized Tourist Agenda. In Proceedings of IADIS International Conference on WWW/INTERNET (Madrid, Spain, October 06-09, 2004). ICWI 2004. IADIS Press, 619-626.

Grabler K. and Zins A. H. (2002). Vacation Trip Decision Styles as Basis for an Automated Recommendation System: Lessons from Observational Studies. In Proceedings of the International Conference on Information Technology and Travel & Tourism. ENTER 2002, 458-469.

Graham A., Garcia-Molina H., Paepcke A., Winograd T. (2002). Time as Essence for Photo Browsing through Personal Digital Libraries. In Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries (Portland, Oregon, USA, July 14-18, 2002). JCDL 2002. ACM Press, New York, NY, USA, 326–335.

Graham M., Kennedy J., Benyon D. (2000). Towards a Methodology for Developing Visualizations. International Journal of Human-Computer Studies, 53: 789-807.

Gretzel U., Hwang Y.-H., Fesenmaier D. R. (2006). A Behavioral Framework for Destination Recommendation Systems Design. Destination Recommendation Systems: Behavioral Foundations and Applications, 53-62.

Grimnes A. A. G., Chalmers S., Edwards P., Preece A. (2003). GraniteNights-A Multi-agent Visit Scheduler Utilizing Semantic Web Technology. In Proceedings of the international workshop on Cooperative Information Agents (Helsinki, Finland, August 27-29, 2003). CIA 2003. Springer Verlag, Berlin/Heidelberg, Germany, 137-151.

Hansen D. L. and Golbeck J. (2009). Mixing It Up: Recommendation Collections of Items. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Boston, USA, April 04-09, 2009). CHI 2009. ACM Press, New York, NY, USA, 1217-1226.

Haralick R. M., Shanmugam K., Dinstein I. H. (1973). Textural Features for Image Classification. IEEE Transactions on Systems, Man, and Cybernetics, 610-621.

Hassan-Montero Y. and Herrero-Solana V. (2006). Improving Tag-Clouds as Visual Music Retrieval Interfaces. In Proceedings of International Conference on Multidisciplinary Information Sciences and Technologies (Mérida, Spain, October 25-28, 2006). InScit 2006.

Hearst M., Elliott A., English J., Sinha R., Swearingen K., Yee K-P. (2002). Finding the Flow in Website Search. Communications of the ACM, 45(9): 42-49.

Hearst M. A. and Rosner D. (2008). Tag Clouds: Data Analysis Tool or Social Signaller? In Proceedings of the 41st Annual Hawaii International Conference on System Sciences (Hawaii, USA, January 7-10, 2008). HICSS 2008. IEEE Computer Society, 1-10.

Heer J. and Boyd D. (2005). Vizster: Visualizing Online Social Networks. In Proceedings of the IEEE symposium on Information Visualization (Minneapolis, MN, USA, October 23-25, 2005). InfoVis 2005. IEEE Computer Society, 33-40.

Heer J., Card S. K., Landay. J. A. (2005). Prefuse: A Toolkit for Interaction Information Visualization. In Proceedings of the SIGCHI conference on Human factors in computing systems (Portland, Oregon, USA, April 2-7, 2005). CHI 2005. ACM Press, New York, NY, USA, 421-430.

Herlocker J. L., Konstan J. A., Riedl J. (2000). Explaining Collaborative Filtering Recommendations. In Proceedings of the ACM conference on Computer supported cooperative work (Philadelphia, PA, USA, December 02-06, 2000). CSCW 2000. ACM Press, New York, NY, USA, 241-250.

Herrera P., Bello J., Widner G., Sandler M., Celma Ò., Vignoli F., Pampalk E., Cano P., Pauws S., Serra X. (2005). SIMAC: Semantic Interaction with Music Audio Contents. In Proceedings of the 2nd European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology (London, UK, November 29-December 01, 2005). EWIMT 2005.

Hertzog P. and Torrens M. (2005). Context-aware Mobile Assistants for Optimal Interaction: A Prototype for Supporting the Business Traveler. In Proceedings of the International Conference on Intelligent User Interfaces (San Diego, CA, USA, January 10-13, 2005). IUI 2005. ACM press, New York, NY, USA, 256-258.

Hilliges O., Baur D., Butz A. (2007). Photohelix: Browsing, Sorting and Sharing Digital Photo Collections. In the Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (Newport, RI, USA, October 10-12, 2007). TABLETOP 2007. IEEE Computer Society, 87–94.

Hilliges O. and Kirk D. S. (2009). Getting Sidetracked: Display Design and Occasioning Photo-Talk with the Photohelix. In proceedings of the SIGCHI conference on Human factors in computing systems (Boston, MA, USA, April 4-9, 2009), CHI 2009. ACM Press, New York, NY, USA, 1733-1736.

Huang Z., Li X., Chen H. (2005). Link Prediction Approach to Collaborative Filtering. In Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries (Denver, CO, USA, June 07-11, 2005). JCDL 2005. ACM Press, New York, NY, USA, 141-142.

Huynh D. F., Drucker S. M., Baudisch P., Wong C. (2006). Time Quilt: Scaling up Zoomable Photo Browsers for Large, Unstructured Photo Collections. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ext. Abstracts (Quebec, Canada, April 22-27, 2006). CHI 2006. ACM Press, New York, NY, USA, 1937–1940.

Hwang Y. H., Gretzelb U., Fesenmaier D. R. (2002). Behavioral Foundations for Human-centric Travel Decision-aid Systems. In Proceedings of the International Conference on Information Technology and Travel & Tourism. ENTER 2002, 356-366.

Jacobsson M., Rost M., Håkansson M., Holmquist L. E. (2005). Push!Music: Intelligent Music Sharing on Mobile Devices. In Adjunct Proceedings of International Conference on Ubiquitous Computing (Tokyo, Japan, September 11-14, 2005). http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.60.924- cited 2010/5/27.

Janiszewski C. (1998). The Influence of Display Characteristics on Visual Exploratory Search Behavior. Journal of Consumer Research, 25(3): 290-301.

Joinson A. N. (2008). “Looking at”, “Looking up” or “Keeping Up With” People?: Motives and Uses of Facebook. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Florence, Italy, April 5-10, 2008). CHI 2008. ACM Press, New York, NY, USA, 1027-1036.

Kang E., Kim H., Cho J. (2006). Personalization Method for Tourist Point of Interest (POI) Recommendation. In Proceedings of the 11th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (Bournemouth, UK, October 09-11, 2006). KES 2006. AMC Press, New York, NY, USA, 392-400.

Kang H. and Shneiderman B. (2000). Visualization Methods for Personal Photo Libraries: Browsing and Searching in the PhotoFinder. In Proceedings of IEEE International Conference on Multimedia & Expo (New York, NY, USA, July 30-August 2, 2000). ICME 2000. IEEE Computer Society, 1539-1542.

Kanungo T., Mount D. M., Netanyahu N. S., Piatko C., Silverman R., Wu A. Y. (2000). Analysis of a Simple k-means Clustering Algorithm. In Proceedings of the sixteenth annual symposium on Computational Geometry (Hong Kong, China, June 12-14, 2000). SoCG 2000. ACM Press, New York, NY, USA, 100-109.

Kaser O. and Lemire D. (2007). Tag-Cloud Drawing: Algorithms for Cloud Visualization. In Proceedings of International World Wide Web Conference (Banff, Alberta, Canada, May 8-12, 2007). WWW 2007. ACM Press, New York, NY, USA.

Keck S. (2009). Incremental Personalized Trip Planning System. Diploma thesis, University of Munich, Germany.

Kennedy L. S., Naaman M., Ahern S., Nair R., Rattenbury T. (2007). How Flickr Helps us Make Sense of the World: Context and Content in Community-Contributed Media Collections. In Proceedings of the 15th international conference on Multimedia (Augsburg, Germany, September 24-29, 2007). MM 2007. ACM Press, New York, NY, USA, 631-640.

Kerr B. (2006). TagOrbitals: Tag Index Visualization. IBM research report, RC24027(W0608-056).

Kim J.-Y. and Berkin N. J. (2002). Categories of Music Description and Search Terms and Phases used by Non-Music Experts. In Proceedings of the International Society for Music Information Retrieval (Paris, France, October 13-17, 2002). ISMIR 2002.

Kirk D., Sellen A., Rother C., Wood K. (2006). Understanding Photowork. In Proceedings of the SIGCHI conference on Human factors in computing systems (Montréal, Québec, Canada, April 22-27, 2006). CHI 2006. ACM Press, New York, NY, USA, 761-770.

Klüber R. (2009). ThumbnailDJ: Visualizing Music Collections based on Visual Thumbnails. Diploma thesis, University of Munich, Germany.

Komninos A., Cameron N., Feron C., Allan R., Lindsay A. (2008). AmbientDJ: Enabling Interaction between People and Socially Aware Environments. In Proceedings of Workshop on Mobile Interaction with the Real Would (Amsterdam, The Netherlands, September 2, 2008). MIRW 2008.

Kroesche J., Baldzer J., Boll S. (2007). MobiDENK-mobile Multimedia in Monument Conservation. IEEE multimedia, 11(2): 72-77.

Kuchinsky A., Pering C., Creech M. L., Freeze D., Serra B., Gwizdka J. (1999). FotoFile: A Consumer Multimedia Organization and Retrieval System. In Proceedings of the SIGCHI conference on Human factors in computing systems (Pittsburgh, PA, USA, April 15-20, 1999).CHI 1999. ACM Press, New York, NY, USA, 496-503.

Kustanowitz J. and Shneiderman B. (2004). Motivating Annotation for Personal Digital Photo Libraries: Lowering Barriers while Raising Incentives. Human-Computer Interaction Lab, University of Maryland, Technical Report HCIL-2004-18.

Kustanowitz J. and Shneiderman B. (2005). Meaningful Presentations of Photo Libraries: Rationale and Applications of Bi-Level Radial Quantum Layouts. In Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries (Denver, Colorado, USA, July 7-11, 2005). JCDL 2005. ACM Press, New York, NY, USA, 188-196.

Kuo B. Y. L., Hentrich T., Good B. M., Wilkinson, M. D. (2007). Tag Clouds for Summarizing Web Search Results. In Proceedings of International World Wide Web Conference (Banff, Alberta, Canada, May 8-12, 2007). WWW 2007. ACM Press, New York, NY, USA, 1203-1204.

Lam H. (2005). MusicLand: Exploratory Browsing in Music Space. Poster in the IEEE symposium on Information Visualization (Minneapolis, MN, USA, October 23-25, 2005). InfoVis 2005.

Lauesen S. (2005). User Interface Design. Pearson/Addison-Wesley, Boston, MA, USA.

Lee J. H. and Downie J. S. (2004). Survey of Music Information Needs, Uses, and Seeking Behaviors: Preliminary Findings. In Proceedings of the International Society for Music Information Retrieval (Barcelona, Spain, October 10-15, 2004). ISMIR 2004.

Leong T. W., Howard S., Vetere F. (2008). Choice: Abdicating or Exercising. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Florence, Italy, April 5-10, 2008). CHI 2008. ACM Press, New York, NY, USA, 715-724.

Li R., Bao S., Yu Y., Fei B., Su Z. (2007). Towards Effective Browsing of Large Scale Social Annotations. In Proceedings of International World Wide Web Conference (Banff, Alberta, Canada, May 8-12, 2007). WWW 2007. ACM Press, New York, NY, USA, 943-952.

Logan B. (2002). Content-based Playlist Generation: Exploratory Experiments. In Proceedings of the International Society for Music Information Retrieval (Paris, France, October 13-17, 2002). ISMIR 2002, 295-296.

Logan B. (2004). Music Recommendation from Song Sets. In Proceedings of the International Society for Music Information Retrieval (Barcelona, Spain, October 10-15, 2004). ISMIR 2004.

Lohmann S., Ziegler J., Tetzlaff L. (2009). Comparison of Tag Cloud Layouts: Task-Related Performance and Visual Exploration. In Proceedings of the 12th international conference on Human-Computer Interaction (Uppsala, Sweden, August 24-28, 2009). INTERACT 2009. Springer Verlag, Berlin/Heidelberg, Germany, 392-404.

Lux M. and Chatzichristofis S. A. (2008). Lire: Lucene Image Retrieval: an Extensible Java CBIR Library. In Proceedings of the 16th ACM International Conference on Multimedia (Vancouver, BC, Canada, October 27-31, 2008). MM 2008. ACM Press, New York, NY, USA, 1085-1088.

Mahato H., Kern D., Holleis P., Schmidt A. (2008). Implicit Personalization of Public Environments Using Bluetooth. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ext. Abstracts (Florence, Italy, April 5-10, 2008). CHI 2008. ACM Press, New York, NY, USA, 3093-3098.

Marchionini G. (2006). Exploratory Search, from Finding to Understanding. Communication of the ACM, 49(4): 41-46.

Marchionini G. and Shneiderman B. (1988). Finding Facts vs. Browsing Knowledge in Hypertext Systems. IEEE Computer, 21(3): 70-80.

Markkula M. and Sormunen E. (2000) End-User Searching Challenges Indexing Practices in the Digital Newspaper Photo Archive. Information Retrieval, 1:259-285.

Marks J., Andalman B., Beardsley P. A., Freeman W., Gibson S., Hodgins J., Kand T., Mirtich B., Pfister H., Ruml W., Ryall K., Seims J., Shieber S. (1997). Design Galleries: A General Approach to Setting Parameters for Computer Graphics and Animation. In Proceedings of the 24th annual conference on Computer graphics and interactive techniques (LA, USA, August 3-8, 1997). SIGGRAPH 1997. ACM Press, New York, NY, USA, 389 - 400.

Marlow C., Naaman M., Boyd D., Davis M. (2006). HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, to read. In Proceedings of the seventeenth conference on Hypertext and hypermedia (Odense, Denmark, August 23-25, 2006). HYPERTEXT 2006. ACM Press, New York, NY, USA, 31-40.

Marshall C. C. and Bly S. (2004). Sharing Encountered Information: Digital Libraries Get a Social Life. In Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries (Tucson, Arizona, USA, July 7-11, 2004). JCDL 2004. ACM Press, New York, NY, USA, 218-227.

McCarthy J. F. and Anagnost T. D. (1998). MusicFX: An Arbiter of Group Preferences for Computer Supported Collaborative Workouts. In Proceedings of ACM Conference on Computer Supported Cooperative Work (Seattle, WA, USA, November 14-18, 1998). CSCW 1998. ACM Press, New York, NY, USA, 363-372.

McEnnis D., McKay C., Fujinaga I., Depalle P. (2005). jAudio: A Feature Extraction Library. In Proceedings of the International Society for Music Information Retrieval (London, UK, October 11-15, 2005). ISMIR 2005.

Mehta B., Hofmann T., Nejdl W. (2007). Robust Collaborative Filtering. In Proceedings of the ACM Conference on Recommender Systems (Minneapolis, MN, USA, October 19-20, 2007). RecSys 2007. ACM Press, New York, NY, USA, 49-56.

Miller A. D. and Edwards W. K. (2007). Give and Take: A Study of Consumer Photo-Sharing Culture and Practice. In proceedings of the SIGCHI conference on Human factors in computing systems (San Jose, California, USA, April 28-May 3, 2007), CHI 2007. ACM Press, New York, NY, USA, 347-356.

Moe W. W. (2007). Buying, Searching or Browsing: Differentiating between Online Shoppers Using In-Store Navigational Clickstream. Journal of Consumer Psychology, 13(1&2):29-39.

Montaner M., López B., de la Rosa J. L. (2003). A Taxonomy of Recommender Agents on the Internet. Artificial Intelligence Review, 19(4): 285-330.

Morris M. R., Paepcke A., Winograd T. (2006a). TeamSearch: Comparing Techniques for Co-Present Collaborative Search of Digital Media. In the Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (Adelaide, South Australia, January 5-7, 2006). TABLETOP 2006. IEEE Computer Society.

Morris M. R., Paepcke A., Winograd T., Stamberger J. (2006b). TeamTag: Exploring Centralized versus Replicated Controls for Co-Located Tabletop Groupware. In Proceedings of the SIGCHI conference on Human factors in computing systems (Montréal, Québec, Canada, April 22-27, 2006).CHI 2006. ACM Press, New York, NY, USA, 1273-1282.

Naaman M., Harada S., Wang Q. Y., Garcia-Molina H., Paepcke A. (2004a). Context Data in Geo-Referenced Digital Photo Collections. In proceedings of the 12th annual ACM international conference on Multimedia (New York, NY, USA, October 10-16, 2004). MM 2004. ACM Press, New York, NY, USA, 196-203.

Naaman M., Paepcke A, Garcia-Molina H. (2003). From Where to What: Metadata Sharing for Digital Photographs with Geographic Coordinates. On The Move to Meaningful Internet Systems. Springer Verlag, Berlin/Heidelberg, Germany, 196-217.

Naaman M., Song Y. J., Paepcke A., Garcia-Molina H. (2004b). Automatic Organization for Digital Photographs with Geographic Coordinates. In Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital libraries (Tucson, Arizona, USA, July 7-11, 2004). JCDL 2004. ACM Press, New York, NY, USA, 53-62.

Naaman M., Yeh R. B., Garcia-Molina H., Paepcke A. (2005). Leveraging Context to Resolve Identity in Photo Albums. In Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries (Denver, Colorado, USA, July 7-11, 2005). JCDL 2005. ACM Press, New York, NY, USA, 178-187.

Nazir A., Raza S., Chuah C. N. (2008). Unveiling Facebook: A Measurement Study of Social Network based Applications. In Proceedings of Internet Measurement Conference (Vouliagmeni, Greece, October 20-22, 2008). IMC 2008. ACM Press, New York, NY, USA, 43-56.

Norman D. A. and Draper S. W. (1986). User Centered System Design: New Perspectives on Human-Computer Interaction. Lawrence Erlbaum Associates, Inc. Mahwah, NJ, USA.

Nov O., Naaman M., Ye C. (2008). What Drives Content Tagging: The Case of Photos on Flickr. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Florence, Italy, April 5-10, 2008). CHI 2008. ACM Press, New York, NY, USA, 1097-1100.

O’Donovan J., Smyth B., Gretarsson B., Bostandjiev S., Höllerer T. (2008). PeerChooser: Visual Interactive Recommendation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Florence, Italy, April 5-10, 2008). CHI 2008. ACM Press, New York, NY, USA, 1085-1088.

O’Hara K., Lipson M., Jansen M., Unger A., Jeffries H., Macer P. (2004). Jukola: Democratic Music Choice in a Public Space. In Proceedings of the IEEE International workshop on Distributed Intelligent Systems (High Tatras, Slovakia, April 14-18, 2004). DIS 2004. 145-154.

O’Hare N., Lee H., Cooray S., Gurrin C., Jones G. J. F., Malobabic J., O’Connor N. E., Smeaton A. F., Uscilowski B. (2006). MediAssist: Using Content-Based Analysis and Context to Manage Personal Photo Collections. In Proceedings of the 5th International

Conference on Image and Video Retrieval (Tempe, AZ, USA, July 13-15). CIVR 2006. Springer Verlag, Berlin / Heidelberg, Germany, 529-532.

Pampalk, E. (2003). Islands of Music: Analysis, Organization, and Visualization of Music Archives. Journal of the Austrian Society for Artificial Intelligence.

Pampalk E. and Goto M. (2006). MusicRainbow: A New User Interface to Discover Artists Using Audio-based Similarity and Web-based Labeling. In Proceedings of the International Society for Music Information Retrieval (Victoria, BC, Canada, October 8-12, 2006). ISMIR 2006.

Pampalk E. and Goto M. (2007). MusicSun: A New Approach to Artist Recommendation. In Proceedings of the International Society for Music Information Retrieval (Vienna, Austria, October 23-30, 2007). ISMIR 2007.

Pampalk E., Pohle T., Widmer G. (2005). Dynamic Playlist Generation based on Skipping Behavior. In Proceedings of the International Society for Music Information Retrieval (London, UK, October 11-15, 2005). ISMIR 2005, 634-637.

Papagelis M., Rousidis I., Plexousakis D, Theoharopoulos E. (2005). Incremental Collaborative Filtering for Highly-Scalable Recommendation Algorithms. In Proceedings of the international symposium on Methodologies for Intelligent Systems (Saratoga Springs, NY, USA, May 25-28, 2005). ISMIS 2005. Springer Verlag, Berlin/Heidelberg, Germany, 553-561.

Pauws S. (2002). PATS: Realization and User Evaluation of an Automatic Playlist Generator. In Proceedings of the International Society for Music Information Retrieval (Paris, France, October 13-17, 2002). ISMIR 2002.

Pauws S. and van de Wijdeven S. (2005). User Evaluation of a New Interactive Playlist Generation Concept. In Proceedings of the International Society for Music Information Retrieval (London, UK, October 11-15, 2005). ISMIR 2005.

Peter S. (2010). TravelDiary - Supporting Storytelling and Sharing of Personal Travel Photo Collections. Diploma thesis, University of Munich, Germany.

Platt J. C., Czerwinski M., Field B. A. (2002). Phototoc: Automatic Clustering for Browsing Personal Photographs. Microsoft Research Technical Report, MSR-TR-2002-17.

Pohle T., Pampalk, E., Widmer G. (2005a). Generating Similarity-based Playlists Using Traveling Salesman Algorithms. In Proceedings of the 7th International Conference on Digital Audio Effects (Madrid, Spain, September 20-22, 2005). DAFx 2005.

Pohle T., Pampalk E., Widmer G. (2005b). Evaluation of Frequently Used Audio Features for Classification of Music into Perceptual Categories. In Proceedings of the Fourth International Workshop on Content-based Multimedia Indexing (Riga, Latvia, June 21-23). CBMI 2005.

Popescu A. and Grefenstette G. (2009). Deducing Trip Related Information From Flickr. In Proceedings of International World Wide Web Conference (Madrid, Spain, April 20-24, 2009). WWW 2009. ACM Press, New York, NY, USA, 1183-1184.

Porter M. F. (1980). An Algorithm for Suffix Stripping. Program 14(3):130-137.

Pousman Z. and Stasko J. (2006). A Taxonomy of Ambient Information Systems: Four Patterns of Design. In Proceedings of the working conference on Advanced Visual Interfaces (Venezia, Italy, May 23-26, 2006). AVI 2006. ACM Press, New York, NY, USA, 67-74.

Pousman Z., Stasko J. T., Mateas M. (2007). Casual Information Visualization: Depictions of Data in Everyday Life. IEEE transaction on Visualization and Computer Graphics, 13(6):1145-1152.

Preece J., Roger Y., Sharp H. (2001). Beyond Interaction Design: Beyond Human-Computer Interaction. Jon Wiley & Sons, Inc. New York, NY, USA.

Pretzlav M. A. (2008). Last.fm Explorer: An Interactive Visualization of Hierarchical Time-Series Data. http://vis.berkeley.edu/courses/cs294-10-fa08/wiki/images/9/9f/PretzlavFinalPaper.pdf-cited 2010/5/27.

Prieur C., Cardon D., Beuscart J. S., Pissard N., Pons P. (2007). The Strength of Weak Cooperation: A Case Study on Flickr. http://arxiv.org/ftp/arxiv/papers/0802/0802.2317.pdf - cited 2010/5/27.

Reiter M. (2009). PhotoMagnets: Supporting Flexible Browsing and Searching within Personal Photo Collections. Diploma thesis, University of Munich, Germany.

Ricci F. and Missier F. (2004). Supporting Travel Decision Making Through Personalized Recommendation. Designing Personalized User Experiences in eCommerce. Springer Verlag, Netherlands, 231-251.

Roberts R. M. (1989). Serendipity: Accidental Discoveries in Science. Jon Wiley & Sons, Inc. New York, NY, USA.

Rodden K., Basalaj W., Sinclair D., Wood K. (2001). Does Organization by Similarity Assist Image Browsing? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Seattle, Washington, USA, March 31-April 5, 2001). CHI 2001. ACM Press, New York, NY, USA, 190-197.

Rodden K. and Wood K. R. (2003). How Do People Manage Their Digital Photographs. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Fort Lauderdale, USA, April 5-10, 2003). CHI 2003. ACM Press, New York, NY, USA, 409-416.

Romero M. and Mateas M. (2005). A Preliminary Investigation of Alien Presence. In Proceedings of the International Conference on Human-Computer Interaction (Las Vegas, Nevada, USA, July 22-27, 2005). HCII 2005. http://www.cc.gatech.edu/~mromero/romero_mateas_2005.pdf- cited 2010/5/27.

Santamaría R. and Therón R. (2008). Overlapping Clustered Graphs: Co-authorship Networks Visualization. In Proceedings of 8th International Symposium on Smart Graphics (Rennes, France, August 27-29, 2008). SG 2008. Springer Verlag, Berlin/Heidelberg, Germany, 190-199.

Saraiya P., North C., Duca K. (2005). An Insight-based Methodology for Evaluating Bioinformatics Visualizations. IEEE transactions on visualization and computer graphics, 11(4):1-14.

Schäfer L, Valle C., Prinz W. (2004). Group storytelling for team awareness and entertainment. In Proceedings of the third Nordic conference on Human-computer interaction (Tampere, Finland, October 23-27, 2004). NordiCHI 2004. ACM Press, New York, NY, USA, 441-444.

Schiel U., Baptista C. S., Maia A. C. J., Andrage F. G. (2007). SEI-Tur: A System based on Composed Web-Service Discovery to Support the Creation of Trip Plans. In Proceedings of the International Conference on Defects in Semiconductors (Albuquerque, New Mexico, USA, July 22-27, 2007). ICDS 2007.

Schmidt M. (2009). MusicTrends: Map-based Visualization of User-Generated Content in Last.fm. Project thesis, University of Munich, Germany.

Schraefel M. C., Smith D. A., Owens A., Russell A., Harris C., Wilson M. L. (2005). The Evolving mSpace Platform: Leveraging the Semantic Web on the Trail of the Memex. In Proceedings of the sixteenth ACM conference on Hypertext and hypermedia (Salzburg, Austria, September 6-9, 2005). HT 2005. ACM Press, New York, NY, USA, 174-183.

Sebastian L., Garcia I., Onaindia E., Guzman C. (2008). e-Tourism: A Tourist Recommendation and Planning Application. In Proceedings of IEEE international conference on Tools with Artificial Intelligence (Dayton, Ohio, USA, November, 03-05, 2008). ICTAI 2008. 89-96.

Seifert C., Kump B., Kienreich W., Granitzer G., Granitzer M. (2008). On the Beauty and Usability of Tag Clouds. In Proceedings of the International Conference Information Visualization (Eindhoven, The Netherlands, June 4-6, 2008). IV 2008. IEEE Computer Society, 17-25.

Sharp H., Rogers Y., Preece J. (2007). Interaction Design: Beyond Human-Computer Interaction. Jon Wiley & Sons, Inc. New York, NY, USA.

Shen C., Lesh N. B., Moghaddam B., Beardsley B., Bardsley P., Bardsley R. S. (2001). Personal Digital Historian: User Interface Design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ext. Abstracts (Seattle, Washington, USA, March 31-April 5, 2001). CHI 2001. ACM Press, New York, NY, USA, 29-30.

Shen C., Lesh N. B., Vernier F., Forlines C., Frost J. (2002). Sharing and Building Digital Group Histories. In proceedings of ACM Conference on Computer Supported Cooperative Work (New Orleans, Louisiana, USA, November 16-20, 2002). CSCW 2002. ACM Press, New York, NY, USA, 324–333.

Shneiderman B. (1992). Tree Visualization with Treemaps: A 2-D Space-Filling Approach. ACM Transactions on Graphics, 11(1): 92–99.

Shneiderman B. (2000). Universal Usability. Commutations of ACM, 43(5):84-91.

Shneiderman B. and Kang H. (2000). Direct Annotation: A Drag-and-Drop Strategy for Labeling Photos. In Proceedings of the IEEE symposium on Information Visualization (Salt Lake City, Utah, USA, October 9-10, 2000). InfoVis 2000. IEEE Computer Society, 88-98.

Shneiderman B. and Plaisant C. (2005). Designing the User Interface. Pearson/Addison-Wesley, Boston, MA, USA.

Sinha R. and Swearingen K. (2002). The Role of Transparency in Recommender Systems. In Proceedings of the International Conference on Human Factors in Computing Systems (Minnesota, USA, April 20-25, 2002). CHI 2002. ACM Press, New York, NY, USA, 830-831.

Skog T., Ljunglblad S., Holmquist L. E. (2003). Between Aesthetics and Utility: Designing Ambient Information Visualization. In Proceedings of the IEEE symposium on Information Visualization (Seattle, Washington, USA, October 19-21, 2003). InfoVis 2003. IEEE Computer Society, 233-240.

Snavely N., Seitz S. M., Szeliski R. (2006). Photo Tourism: Exploring Photo Collections in 3D. In Proceedings of the International Conference on Computer Graphics and Interactive Techniques (Boston, Massachusetts, USA, July 30-August 3, 2006). SIGGRAPH 2006. ACM Press, New York, NY, USA, 835-846.

Spence R. (2001). Information Visualization. Pearson/Addison-Wesley, Boston, MA, USA.

Stasko K., Miller T., Pousman Z., Plaue C., Ullah O. (2004). Personalized Peripheral Information Awareness through Information Art. In Proceedings of International Conference on Ubiquitous Computing (Nottingham, UK, September 7-10, 2004). UbiComp 2004. Springer Verlag, Berlin/Heidelberg, Germany, 18-35.

Stasko J. and Zhang E. (2000). Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations. In Proceedings of the IEEE symposium on Information Visualization (Salt Lake City, Utah, USA, October 9-10, 2000). InfoVis 2000. IEEE Computer Society, 57.

Stavness I., Gluck J., Vilhan L., Fels S. (2005). The MUSICtable: A Map-based Ubiquitous System for Social Interaction with a Digital Music. Lecture Notes in Computer Science. Springer Verlag, Berlin/Heidelberg, Germany, 291-302.

Stricker M. and Orengo M. (1995). Similarity of Color Images. In Proceedings of Storage and Retrieval for Image and Video Databases (San Diego, CA, USA, July 9, 1995). SPIE 1995. 381–392.

Swearingen K. and Sinha R. (2001). Beyond Algorithms: An HCI Perspective on Recommender Systems. In ACM SIGIR Workshop on Recommender Systems (2001), ACM Press, New York, NY, USA, 1-11.

Swearingen K. and Sinha R. (2002). Interaction Design for Recommender Systems. In Proceedings of the Conference on Designing Interactive Systems (London, UK, June 25-28, 2002). DIS 2002.

Toms E. G. (2000a). Serendipitous Information Retrieval. In Proceedings of the First DELOS Network of Excellence Workshop on Information Seeking, Searching and Querying in Digital Libraries.
http://www.ercim.org/publication/ws-proceedings/DelNoe01/3_Toms.pdf- cited 2010/5/27.

Toms E. G. (2000b). Understanding and Facilitating the Browsing of Electronic Text. International Journal of Human Computer Studies, 52(3): 423-451.

Torrens M., Hertzog P., Arcos J. L. (2004). Visualizing and Exploring Personal Music Libraries. In Proceedings of the International Society for Music Information Retrieval (Barcelona, Spain, October 10-15, 2004). ISMIR 2004.

Torres R. S., Silva C. G., Medeiros C. B., Rocha H. V. (2003). Visual Structures for Image Browsing. In Proceedings of the twelfth international conference on Information and knowledge management (New Orleans, LA, USA, November 2-8, 2003). CIKM 2003. ACM Press, New York, NY, USA, 49-55.

Toyama K., Logan R., Roseway A. (2003). Geographic Location Tags on Digital Images. In Proceedings of the 11th international conference on Multimedia (Berkeley, CA, USA, November 2-8, 2003). MM 2003. ACM Press, New York, NY, USA, 156-166.

Van Gulik R. and Vignoli F. (2005). Visual Playlist Generation on the Artist Map. In Proceedings of the International Society for Music Information Retrieval (London, UK, October 11-15, 2005). ISMIR 2005.

Van House N., Davis M., Ames M., Finn M., Viswanathan V. (2005). The Uses of Personal Networked Digital Imaging: An Empirical Study of Cameraphone Photos and Sharing. In Proceedings of the SIGCHI conference on Human factors in computing systems, Ext. Abstracts (Portland, Oregon, USA, April 2-7, 2005).CHI 2005. ACM Press, New York, NY, USA, 1853-1856

Van House N., Davis M., Takhteyev Y., Good N., Wilhelm A., Finn M. (2004). From What to Why: Social Use of Personal Photos. In Proceedings of the 2004 ACM conference on Computer Supported Cooperative Work, Ext. Abstracts (Chicago, Illinois, November 6-10, 2004). CSCW 2004. ACM Press, New York, NY, USA.

Vickery, J. E. (1997). A Note in Defense of Browsing. BLL review, 5:110.

Viégas F. B., Golder S., Donath J. (2006). Visualizing Email Content, Portraying Relationships from Conversational Histories. In Proceedings of the SIGCHI conference on Human factors in computing systems (Montréal, Québec, Canada, April 22-27, 2006). CHI 2006. ACM Press, New York, NY, USA, 979-988.

Viégas F. B., Perry E., Howe E., Donath J. (2004). Artifacts of the Presence Era: Using Information Visualization to Create an Evocative Souvenir. In Proceedings of the IEEE symposium on Information Visualization (Austin, Texas, USA, October 10-12, 2004). InfoVis 2004. IEEE Computer Society, 105-111.

Viégas F. B. and Wattenberg M. (2007). Artistic Data Visualization, beyond Visual Analytics. In Proceedings of the International Conference on Human-Computer Interaction (Beijing, China, July 22-27, 2007). HCII 2007. Springer Verlag, Berlin/Heidelberg, Germany, 182-191.

Viégas F. B. and Wattenberg M. (2008). Tag Clouds and the Case for Vernacular Visualization. Interactions, 15(4): 49-52.

Vignoli F. (2004). Digital Music Interaction Concepts: A User Study. In Proceedings of the International Society for Music Information Retrieval (Barcelona, Spain, October 10-15, 2004). ISMIR 2004.

Vignoli F. and Pauws S. (2005). A Music Retrieval System based on User-Driven Similarity and Its Evaluation. In Proceedings of the International Society for Music Information Retrieval (London, UK, October 11-15, 2005). ISMIR 2005.

Voida A., Grinter R. E., Ducheneaut N., Edwards W. K., Newman M. W. (2005). Listening in: Practices Surrounding iTunes Music Sharing. In Proceedings of the SIGCHI conference on Human factors in computing systems (Portland, Oregon, USA, April 2-7, 2005). CHI 2005. ACM Press, New York, NY, USA, 191-200.

Von Ahn. L. and Dabbish L. (2006). Labeling Images with a Computer Game. In Proceedings of the SIGCHI conference on Human factors in computing systems (Portland, Oregon, USA, April 24-29, 2006).CHI 2006. ACM Press, New York, NY, USA, 319-326.

Ware C. (2004). Information Visualization: Perception for Design. Morgan Kaufmann, San Fransisco, CA, USA.

White R. W., Kules B., Drucker S. M. (2006). Supporting Exploratory Search. Communications of the ACM, 49(4): 37-39.

Wijk J. V. (2006). Bridging the Gaps. IEEE Computer Graphics and Applications, 6-9.

Williamson C. and Shneiderman B. (1992). The Dynamic HomeFinder: Evaluating Dynamic Queries in a Real-Estate Information Exploration System. In Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval (Copenhagen, Denmark, June 21-24, 1992). SIGIR 1992. ACM Press, New York, NY, USA, 338-346.

Wilson M. L. (2007). A Nine Month Report on Progress towards a Framework for Evaluating Advanced Search Interfaces Considering IR and HCI. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.2797&rep=rep1&type=pdf - cited 2010/5/27.

Wohltorf J., Cissée R., Rieger A. (2005). BerlinTainment: An Agent-based Context-Aware Entertainment Planning System. IEEE Communications, 43(6):102-10.

Wu X., Zhang L., Yu Y. (2006). Exploring Social Annotations for the Semantic Web. In Proceedings of International World Wide Web Conference (Edinburgh, UK, May 22-26, 2006). WWW 2006. ACM Press, New York, NY, USA, 417-426.

Xiong R. and Donath J. (1999). PeopleGarden: Creating data Portraits for Users. In Proceedings of the 14th annual ACM Symposium on User interface Software and Technology (Asheville, North Carolina, United States, November 7-10, 1999). UIST 1999. ACM Press, New York, NY, USA, 37-44.

Yanai K. and Qiu B. (2009). Mining Cultural Differences from a Large Number of Geotagged Photos. In Proceedings of International World Wide Web Conference (Marid, Spain, April 20-24, 2009). WWW 2009. ACM Press, New York, NY, USA, 1173-1174.

Yee K. P., Swearingen K., Li K., Hearst M. (2003). Faceted Metadata for Image Search and Browsing. In Proceedings of the SIGCHI conference on Human factors in computing systems (Fort Lauderdale, Florida, USA, April 5-10, 2003). CHI 2003. ACM Press, New York, NY, USA, 401-408.

Yi J. S., Melton R., Stasko J., Jacko J. A. (2005). Dust & Magnet: Multivariate Information Visualization Using a Magnet Metaphor. In Proceedings of the IEEE symposium on Information Visualization (Minneapolis, MN, USA, October 23-25, 2005). InfoVis 2005. IEEE Computer Society, 239-256.

Zhou X. S. and Huang T. S. (2003). Relevance Feedback in Image Retrieval: A Comprehensive Review. Special Issues of the journal of Multimedia Systems on Content-based Image Retrieval. Springer Verlag, Berlin/Heidelberg, Germany, 8(6): 536-544.