TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

Arto Lehtiniemi

Novel Music Discovery Concepts: User Experience and Design Implications

Julkaisu 1257 • Publication 1257

Tampere 2014

Tampereen teknillinen yliopisto. Julkaisu 1257
Tampere University of Technology. Publication 1257

Arto Lehtiniemi

Novel Music Discovery Concepts: User Experience and Design Implications

Thesis for the degree of Doctor of Science in Technology to be presented with due permission for public examination and criticism in Tietotalo Building, Auditorium TB109, at Tampere University of Technology, on the 28th of November 2014, at 12 noon.

Tampereen teknillinen yliopisto - Tampere University of Technology
Tampere 2014

ISBN 978-952-15-3392-1 (printed)
ISBN 978-952-15-3432-4 (PDF)
ISSN 1459-2045

Abstract

Current music consumers are facing an almost endless selection of music in online services to be accessed on-demand with a variety of devices. The focus has now shifted from providing on-demand access to massive music catalogs towards improving the user experience of the music services, providing new ways of finding relevant music from the massive online catalogs, and making music consumption a pleasurable experience. The key differentiation aspects for music services come largely from the user interface and the ways that music can be found or consumed.

This thesis belongs to the fields of human-computer interation (HCI) and music information retrieval (MIR). HCI is concerned with the design, evaluation and implementation of interactive computing systems and MIR targets to broaden the understanding and usage of musical data through research, applications and tools.

This thesis studies novel concepts for music discovery that are based on strong visual metaphors and stereotypes. The goal is to research the user experience (UX) of novel music discovery services and to formulate key design implications to support service development for music discovery. The research of music discovery prototypes consisted of three main phases: initial concept design phase, playful concept exploration phase, and iterative concept design phase. The thesis introduces, in total, ten prototype implementations of these novel concepts for music discovery. User evaluations of the implemented prototypes were conducted with Finnish active music listeners with both qualitative and quantitative research methods. This thesis contributes to both academic research on HCI in MIR and commercial music discovery service development.

The results provide insights to user experience with different types of novel music discovery services. Five novel music discovery services using the same content-based music recommendation back-end were compared and the comparison results are reported including both first impressions and longer-term usage. Additionally, the results of the studies introduce a wide set of future directions for each music discovery approach. These future directions enable service developers to further enhance the music discovery experience within these fields.

All but one of the proposed music discovery concepts work well for music discovery. The use of avatar characters and mood pictures for music discovery are the most promising ones. The results show that visual music discovery services have the potential to replace traditional music discovery services in different types of music discovery practices.

The final contribution of the thesis is a set of 16 design implications for music discovery service developers.

Preface

This work has been carried out at Nokia Research Center (NRC) from 2007 to 2014. First and foremost, I would like to thank Professor Kaisa Väänänen-Vainio-Mattila (TUT) for supervising my work and the encouragement and support of the research in the crossroads of HCI and MIR. Her active guidance and positive attitude towards pursuing my ambitions in the field of music discovery has made a significant impact to the work. I would like to thank the pre-examiners of this thesis, Professor Perfecto Herrera and Associate Professor Catherine Guastavino for the insightful comments and Sergi Jordà for being my opponent at the public examination.

The thesis would not been possible without the significant support from my current and previous Nokia managers and colleagues. First, I wish to thank Mauri Väänänen for believing in me and letting me to steer my audio-related work towards the field of usability and user experience already in 2001, when the rest of the team was doing audio coding. It was a bold move, Mauri! Next, I want to thank the incredible SuperMusic team: Jarno Seppänen, Antti Eronen, Jukka Holm, Mikko Heikkinen, and Marko Takanen. We really did something superior at the time! I want to thank Kari Laurila for managing the SuperMusic team and letting people believe in their intuition in order to achieve something special. Kari's impact to my thesis was still even bigger. I want to express my gratitude for him in letting me pursue my ambition in the domain of visual music discovery concepts. He enabled me to design and develop several prototypes featured in this thesis. Special thanks go to Jukka Holm for taking the task of evaluating these prototypes with me and co-authoring several research papers featured in this thesis. Jukka has been a great colleague and friend to work with and I look forward to researching the domain of music discovery with him even deeper in the future!

When the focus of my main work activities at Nokia steered towards new areas, Jari Hämäläinen, Guido Grassel and Ville-Veikko Mattila played a significant role in still supporting the continuation of music discovery work as a side activity regardless of the internal situation at the company. Additionally, Ville-Veikko pushed me to finalize the thesis and offered great support. Thanks guys! I also want to thank Jyri Huopaniemi for his support along the way as head of laboratory.

I wish to thank Jukka Saarinen for his contribution to the Demola activities and Demola together with all participated students for the prototype implementations. I wish to thank Jarno Ojala (TUT) for helping me to evaluate several prototypes and co-authoring research papers.

Next, I want to thank the music analysis gurus Antti Eronen and Jussi Leppänen for long discussions on various music and non-music related topics, providing back-end solutions for the prototypes, support in MIR-domain of the thesis and related publications, and achieving incredible results together in the field of audio analysis!

Furthermore, I want to thank Miikka Vilermo and David Isherwood for proofreading my thesis and various research papers. I highly appreciate the fruitful discussions with Miikka regarding music discovery and music in general. I also want to thank Nokia Foundation for a scholarship in 2013 to finalize my thesis.

Finally, my biggest thanks go to my ladies at home (Katja, Anni and Maija) for being there for me and the support and encouragement to write the thesis.

Tampere 15.8.2014 Arto Lehtiniemi

Supervisor: Professor Kaisa Väänänen-Vainio-Mattila
Department of Software Systems - Human-Centered Technology
Tampere University of Technology

Pre-examiners: Professor Perfecto Herrera
Music Technology Group
Universitat Pompeu Fabra, Barcelona, Spain

Associate Professor Catherine Guastavino
McGill University, Montreal, Canada

Opponents: Associate Professor Sergi Jordà
Music Technology Group
Universitat Pompeu Fabra, Barcelona, Spain

Contents

List of publications

The thesis consists of a summary and the following original publications:

P1 Lehtiniemi, A. (2008). Evaluating SuperMusic: streaming context-aware mobile music service. *Proceedings of the 2008 International Conference on Advances in Computer Entertainment* (pp. 314-321). ACM Press.

P2 Holm, J., Lehtiniemi A. (2011). A virtual world prototype for interacting with a music collection. *Proceedings of the 4th International Conference on Online Communities and social Computing, held as Part of HCI International 2011* (pp. 326-335). Springer-Verlag.

P3 Holm, J., Lehtiniemi A., & Eronen, A. (2010). Evaluating an avatar-based user interface for discovering new music. *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia.* ACM Press.

P4 Lehtiniemi, A., Holm, J. (2011). Evaluating a potentiometer-based graphical user interface for discovering new music. *Proceedings of the 15th International Conference on Information Visualisation* (pp. 110-118). IEEE Computer Society.

P5 Lehtiniemi, A., Holm, J. (2011). Evaluating a cube-based user interface for exploring music collections. *Proceedings of the 15th International Conference on Information Visualisation* (pp. 40-46). IEEE Computer Society.

P6 Lehtiniemi, A., Holm, J. (2011). Easy access to recommendation playlists: selecting music by exploring preview clips in album cover space. *Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia* (pp. 94-99). ACM Press.

P7 Lehtiniemi, A., Holm, J. (2012). Using animated mood pictures in music recommendation. *Proceedings of the 16th International Conference on Information Visualisation* (pp. 143-150). IEEE Computer Society.

P8 Lehtiniemi, A., Holm, J. (2013). Designing for music discovery: evaluation and comparison of five music player prototypes. *Journal of New Music Research, 42(3),* 283-302. Routledge.

P9 Lehtiniemi, A., Ojala, J. (2012). MyTerritory – evaluation of outdoor gaming prototype for music discovery. *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia.* ACM Press.

P10 Lehtiniemi, A., Ojala, J. (2013). Evaluating MoodPic – a concept for collaborative mood music playlist creation. *Proceedings of the 17th International Conference on Information Visualisation* (pp. 86-95). IEEE Computer Society.

P11 Lehtiniemi, A., Ojala, J. (2014). Using adaptive avatars for visualizing recent music listening history and supporting music discovery. Accepted to *the 11th International Conference on Advances in Computer Entertainment Technology, 11-14 November, Funchal, Madeira.*

1. Introduction

This chapter first describes the background and motivation of the thesis and outlines the scope of the research. This is followed by introducing the research questions and and giving an overview of the included publications. Finally, the key contributions of the thesis are presented.

1.1 Background and motivation

Music is one of the most enjoyed experiences for many people and plays a significant role in modern society regardless of age, gender or culture. Music can be experienced consciously by listening actively to a song, playing an instrument, or listening to music in the background while doing something else. Unconscious exposure to music and sounds can happen almost anywhere, e.g., in shopping centers and while commuting. An anechoic chamber is probably one of the few places where a person is not exposed to any external sound sources.

Music can be used for various purposes such as general enjoyment, mood enhancement, sports, relaxation, and even therapy. In Juslin & Sloboda (2001), the sociological studies reveal that people are not only passively being affected by music, but actively selecting appropriate music to be moved and to seek for a specific change in emotions. Users seem to be well aware of the type of music to be played in different situations for influencing aspects of themselves. Music can be listened to alone or enjoyed together with others. There are many ways to select music to be listened to. Music selection can be highly focused to a specific song or artist, or be more exploratory without a clear goal in mind.

The music industry has faced many changes over time regarding different formats for music consumption. During the last decade, the industry has faced one of the most significant changes yet with the transition from physical CDs towards digital distribution of music. One of the turning points was the introduction of Napster in 1999, after which the sales of physical albums started to decline rapidly. In 2013, 39% of the global music industry revenue (equivalent to almost 6 billion US$) came from digital channels and the numbers are constantly increasing (IFPI 2014). The music business continues to expand into new markets and explore new business models to attract even more users and to bring music to wider audiences.

Digital music is slowly changing the existing ownership model of music to a new model based on subscription and access to music. In 2013, major subscription-based music services offered instant access to up to 37 million songs. Subscription services (including free and paid) grew by over 51% in 2013 and the amount of payed subscribers increased from 8 to 28 million from 2011-2013 (IFPI 2014). Music service providers are actively seeking ways to get even more users to shift from commonly advertiser-funded and restricted streaming options to paid subscriptions.

Music subscription services such as Spotify (2006) have changed the way people search, pay, and experience music. The ability to dynamically create and modify playlists in real time has had a significant impact to our music listening. Spotify, as an example, is growing rapidly and entered 38 new markets in 2013, now being present in a total of 55 markets. There are currently over 1.5 billion ready-made playlists available in Spotify (IFPI 2014).

There is an almost endless selection of music available for consumers and the focus has now shifted from enabling instant access to massive music catalogs towards improving the user experience of the music services, providing new ways to find relevant music from online catalogs, and making music consumption a pleasurable experience. Since a massive amount of music is available from many music services, the key differentiation comes largely from the user interface and the ways the music can be discovered, accessed and consumed. A model of '*satisficing*' was introduced by Bentley et al. (2006). People have an undefined need to listen to music, but may not have an exact idea what to listen to. They choose a set of songs to start with and eventually change their minds, get bored, and skip songs. Eventually they '*satisfice*' to music that matches their current preference well enough. This describes well a common music discovery scenario and opens a window for the development of novel solutions to support the searching and browsing habits.

Hoashi et al. (2009) state that a well-designed interface can improve the usability and efficiency of a system. This is complemented by Norman (2004, p. 2) with "*attractive things work better*". Attractive things have the ability to make people feel good, which enhances creative thinking, helping in problem solving. Additionally, people in a positive mental state are more effective in finding alternative solutions for minor difficulties with a user interface. This thesis focuses on the user experience (UX) of novel music discovery concepts following the ISO-9241-210 (2010) definition: "*person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service*". Aesthetics, playfulness, pleasure, fun, and stimulation play an integral role in overall user experience (Hassenzahl 2008, Korhonen et al. 2009).

Many music services are built using traditional textual lists complemented by album cover art and possibly an artist photograph when displaying the artist biography (if available). There are plenty of opportunities for using more visual and intuitive approaches for music discovery (i.e., finding appropriate and/or new music for the listener). The new exploration opportunities are well described by Schoenwald & Lead (2012, p. 8): *"Five hundred items in a list feels like a lot of work to go through, but presenting a portion of that* *same list in a nonlinear landscape becomes an intriguing exploration*.” The limitations are often related to the amount of descriptive metadata to be used as bases for visual elements. Obtaining new accurate and rich metadata will open new opportunities for user interface and interaction design.

One opportunity is to simplify the design, focusing solely on music discovery providing an entertaining experience for ordinary people without extensive musical knowledge. This type of highly focused approach could be used to complement existing ways to consume music. A potentially useful approach is to build a music service from the visual music discovery viewpoint, still at the same time providing existing means to consume music.

1.2 Scope of the thesis

This thesis studies a variety of novel concepts and their prototypes for visual music discovery. The goal is to research the user experience (UX) of new types of music discovery services and find out key design implications to support the service development for music discovery. The research was conducted at the Nokia Research Center (NRC) in an industrial research context.

The topic is related to the fields of human-computer interation (HCI) and music information retrieval (MIR).

Hewett et al. (1992, Section 2.1) define human-computer interaction as “*a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them*”. Karray et al. (2008, p. 138) further describe human-computer interaction as a “*design that should produce a fit between the user, the machine and the required services in order to achieve a certain performance both in quality and optimality of the services*”.

Music Information Retrieval (MIR) is an interdisciplinary research field targeting to broaden the understanding and usage of musical data through research, applications and tools using knowledge from computer science, music, signal processing and cognition (Casey et al. 2008).

This thesis covers music discovery applications from a HCI perspective and will not cover deep technical details of the researched concepts and their prototypes. The scope of the research is the overall user experience and users’ perceptions towards visual ways to discover music. The proposed solutions aim to simplify music discovery and bring new music to a broader range of users in an entertaining, fun, and playful way regardless of their prior musical knowledge.

1.3 Research questions

This thesis addresses the following research questions:

RQ1. What is the user experience of novel music discovery prototypes?

RQ2. What kind of playful concepts work well for music discovery?

RQ3. What are the design implications for music service developers?

RQ1 investigates the user experience of the described novel concepts and the implemented prototypes separately. This thesis uses the ISO 9241-210 (2010) definition for user experience: "*person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service*". Through a series of user studies, the goal was to learn what type of different (positive and negative) experiences are encountered while using a variety of different types of music discovery systems. A music discovery prototype or system is defined in the thesis to include an application or service that has a major focus on supporting the user in finding new music.

For RQ2, the user evaluation results are analysed as a whole. The results show which playful concepts worked well for music discovery and what are the prominent features of such playful music discovery concepts. Well working concepts in this thesis refer to the positive user experiences in the context of music discovery, based on the following measured attributes: ease of use, fun, suitability for music discovery, and the overall assessment of the concept. Korhonen et al. (2009) define playful experiences through video games such that: "*the behavior is pleasurable, rewarding or reinforcing and that the behavior is exaggerated, are important characteristics of playful experiences*." This thesis investigates playfulness as a characteristic of a system fulfilling several categories from the playful experiences (PLEX) model (Korhonen et al. 2009): captivation, humor, exploration, and discovery (further explained in Section 2.5.3). Additionally, a playful concept in this thesis is considered to be fun and entertaining.

RQ3 takes the findings from all the prototype evaluations and formulates general design implications for music discovery services. These implications help both academics and developers in the research and design of such systems. The formulated list of design implications can be used from the early concepting phase to evaluating existing systems.

All the included publications (and user studies) contribute to all the research questions, excluding **P1** contribution to RQ2 due to not considered being playful.

1.4 Overview of the publications and research process

This dissertation summarizes prototype evaluation results from eleven publications (ten conference papers and one journal paper). Figure 1 illustrates the three phased iterative research process in terms of included publications. Each concept includes a qualitative and quantitative user study to research the user experience.

In order to form valid design implications for music discovery services, the research needed to address many different solutions. Each publication brings new aspects to the design implications and forms an overall view of user experience towards novel music discovery concepts. Additionally, the iterative nature of the research process in the search for the ultimate music discovery experience resulted in a large number of implemented prototypes.

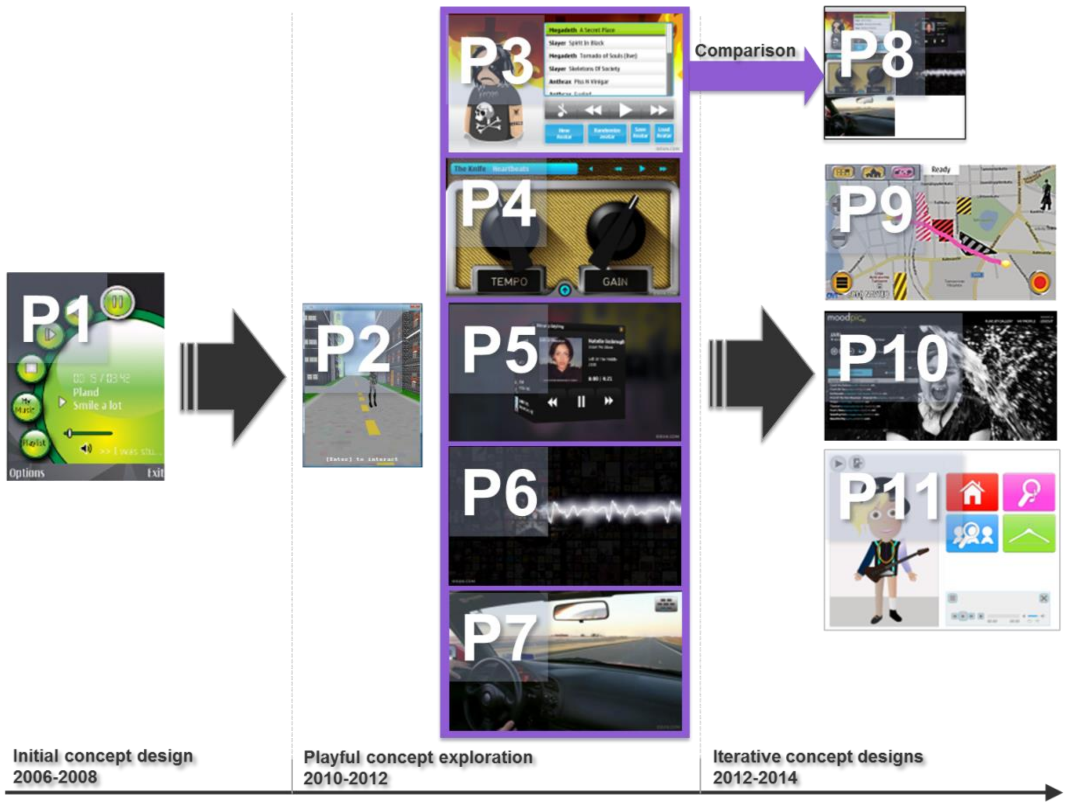

Figure 1 The iterative research process in this thesis is divided into three main phases: Initial concept design, playful concept exploration, and iterative concept design.

The research on music discovery started with the conceptualisation of one of the first mobile streaming music discovery systems, SuperMusic. SuperMusic was developed and evaluated in Nokia Research Center in 2006-2008 (**P1**). The aim was to understand the characteristics of innovative music discovery methods. **P1** introduces a mobile music streaming service with content-based music recommendations and an ability to share music between friends (one-to-one and one-to-many) in the system. Based on the evaluation results, the service was well received and several interesting further development directions were discovered: finding relevant music from large music catalogs, improving music recommendation quality and emphasizing social interaction within a music service. Providing easy and entertaining access to large music catalogs steered the research towards developing visual and playful music discovery concepts in the search for the ultimate music discovery experience. Several new prototypes in publications (**P2-P7**) were developed in the playful concept exploration phase during 2010-2012 (Figure 1) to research different visual ways for accessing large music collections without the need for extensive musical knowledge. These prototypes in publications (**P2-P7**) were developed simultaneously and evaluated in one large user study.

Publication **P2** evaluates a virtual world prototype for music discovery. In the prototype, the user is able to explore a small 3D virtual world city and encounter different looking characters. These characters recommend new music based on their appearance and users are able to add the recommended songs to their playlist. Music can be listened to while exploring and interacting with the virtual world. The qualitative and quantitative evaluation results found the concept not fitting well to users' music consumption habits making the music discovery too time consuming, difficult, and boring.

Publication **P3** introduces a playful prototype for discovering new music using avatar characters. The automatically generated playlist is seeded with three avatar parts: head, torso and the background image. The user is able to change these parts and experiment with different avatar appearances for creating the playlist. Based on the qualitative and quantitative evaluation results the concept inspired users to explore new music and provided fast access to new playlists. In longer-term use, the prototype was found slightly too simple and a number of improvement ideas were introduced.

Publication **P4** studied the idea of using potentiometers for music discovery. In the prototype, the user can first select a preset genre that changes the appearance (i.e., theme) of the prototype. The user can then adjust big tempo and energy level potentiometers to discover music within that genre. While adjusting the potentiometers, the user is able to hear preview clips of representative music. The resulting playlist can be generated based on the potentiometer settings. The qualitative and quantitative evaluation results found the prototype well suited for music discovery, easy to use, and entertaining. In longer-term use the prototype was found slightly too simple.

Publication **P5** investigates the use of a cube metaphor for music exploration. In the prototype, the selected genres are visually built out of small cubes. Each cube can be selected and maximized. A playlist from that genre, music player controls and related information are mapped on the faces of the cube. Qualitative and quantitative evaluation results show potential in the concept and it was found to be an interesting new alternative for ordinary playlists and music discovery. In longer-term use the prototype was found slightly too simple.

Publication **P6** studies music discovery by exploring preview clips in a space filled with album covers. In the prototype, the user is able to move around the abum cover space by moving their finger and corresponding preview clips of different musical styles are played back as a result. Once the user finds suitable music, the finger can be released and a new playlist is generated. The qualitative and quantitative evaluation results show that the concept was seen to be innovative and a fast way for discovering new music without the need for a deeper understanding of music. The instant feedback for finger movement in terms of playing preview clips was appreciated. In longer-term use the prototype was found slightly too simple, although the results show slight increase in the overall experience during the usage period.

Publication **P7** introduces animated mood pictures for music discovery. In the prototype the user is able to select a preset picture that best matches their current or desired mood. A playlist is automatically generated based on the selection and the user is able to further modify and extend the playlist with additional controls. The qualitative and quantitative evaluation results show mood pictures as a promising way of exploring music collections and discovering music. In longer-term use the prototype was criticized for not providing the ability to personalize the pictures and associated music.

In addition to the individual evaluation results of **P2-P7**, related comparison user evaluation data was combined, processed, and published in **P8**. The comparison in **P8** includes five prototypes (**P3-P7**, purple background in Figure 1) since they share the same music recommendation algorithm back-end and music catalog and thus enable shifting focus towards the actual concepts, interaction, and user interface design instead of differences in the technical implementation. Based on the qualitative and quantitative results, the overall winner of the comparison was **P6** (Album cover space). It was described to be closest to a ready-made concept for serious music discovery and performed better than others in longer-term use. Based on the findings the paper introduces design implications for music service developers.

While working on the comparison journal paper, the individual findings from **P2-P7** inspired researching a new domain of music discovery in outdoor environments and started the iterative concept design phase by producing yet another prototype (**P9**).

Publication **P9** introduces a gaming prototype called MyTerritory. The prototype enables users to conquer physical areas by circulating them. The conqueror is able to assign a song to the territory. New areas can be conquered and existing areas can be captured by circulating them several times. The assigned tracks on each territory eventually form a playlist and the users are able to listen to different territories. The qualitative and quantitative results show that the concept adds novel experiences to outdoor gaming and music discovery. The competition in a game was seen as a motivating way to discover and share music by many users. The paper proposes design implications for music service developers.

Despite the overall winner of the prototype comparison in **P8**, the results show more future potential in other concepts (Avatars and mood pictures, **P3** and **P7**) with certain fundamental improvements. Users liked the use of mood pictures and avatars for music discovery but the implementations were lacking important features, e.g., having too little user control over the playlists (**P3**) and customizing pictures and music associations (**P7**). These two (**P3** and **P7**) concepts received a larger amount of further development ideas to enhance the user experience. So far, the prototypes **P2-P7** are preferred mostly to complement other music software for discovery and do not fulfill all of the music consumption requirements. The iterative concept design phase continued with two new prototypes, **P10** and **P11** (Collaborative mood pictures and adaptive avatars). The new concepts are based on **P3** and **P7** but are greatly improved based on the earlier findings.

Publication **P10** further studies the use of mood pictures for music discovery. The prototype enables the users to upload their own pictures as a basis for playlists or to use the newly designed preset mood pictures. All the playlists in the system can be collaboratively populated by adding new songs. The designed preset mood pictures with playlists are provided for easy access to music matching common moods. The qualitative and quantitative results show that accessing music through mood pictures was highly appreciated and was seen as a good way for music discovery. Accessing music through mood pictures instead of genres was seen as an interesting way to interact with music. The paper proposes design implications for music service developers.

Publication **P11** studies the use of adaptive avatars for visualizing recent music listening history and supporting music discovery. The prototype introduces an initially naked avatar character that rapidly reflects the music consumption of the user by populating the avatar parts with matching clothing and accessories. The avatar parts associate information with the songs that triggered them. The user is able to listen to the songs associated with individual avatar parts as well as the whole avatar. Users are able to explore and listen to other users' avatars and recommend avatar parts as visual recommendations to other users. The qualitative and quantitative results show the concept to be fun and providing new music listening and discovery experiences. Users were highly interested in the constantly evolving avatar and exploring other users' avatars. The users had different avatar creation strategies that well support music discovery. Based on these findings the paper proposes design implications for music service developers.

The researched prototypes are summarized below:

- SuperMusic (**P1**): streaming mobile music service with personal and content-based recommendations
- Virtual world (**P2**): music discovery in a 3D virtual world
- Avatar (**P3**): creating playlists through avatar manipulation
- Potentiometers (**P4**): selecting a genre theme and fine-tuning the tempo and gain potentiometers to retrieve a playlist
- Cube (**P5**): mapping musical information to the faces of a cube was used for music discovery
- Album cover space (**P6**): moving a finger over album covers for different types of preview clips of songs to be used for creating a playlist
- Mood pictures (**P7**): preset mood pictures for automatic playlist generation and music discovery
- MyTerritory (**P9**): circulate and conquer physical territories from a map interface to associate songs and discover music from these territories
- Collaborative mood pictures (**P10**): build mood playlists collaboratively using preset or user-generated pictures
- Adaptive Avatar (**P11**): discover music using avatar characters visualizing recent music listening history

In **P1** the author was part of the SuperMusic design and concepting team, designed the prototype user interface and the user trials. The author implemented the user trials, processed the results and wrote the paper. In **P2-P3** the author was a co-inventor of the concepts, led the projects in creating the prototypes, designed and implemented the user study together with Jukka Holm, processed and analyzed the qualitative results and focused on the prototype concepts and future directions. In **P4-P8** the author was a co-inventor of the concepts, led the project in creating the prototypes, designed and implemented the user study together with Jukka Holm, processed and analyzed the qualitative results and focused on the prototype concept, future directions, and design implications. In **P9-P11** the author was a co-inventor or sole inventor of the concepts, led the projects in creating the prototypes, designed the user study together with Jarno Ojala, and wrote the majority of the papers.

1.5 Contributions

This thesis contributes to the work of both academic researchers and commercial music discovery service developers. The focus is on Western geographical music markets and users, as the studies have been conducted in such context.

The thesis introduces ten different approaches for music discovery with implemented prototypes and related user evaluations: streaming mobile recommendation system, virtual world, avatars, potentiometers, cube, album cover space, animated mood pictures, outdoor music discovery, collaborative mood pictures, and adaptive avatars. The user evaluation results are reported in this thesis to give an overall picture of how active listeners of music in Western geographical markets perceive different types of novel music discovery services. Five novel music discovery services using the same content-based music recommendation back-end are compared and the comparison results are reported, including first impressions and longer-term usage. A similar comparison study for visual music discovery systems does not exist. Additionally, the studies introduce a wide set of future directions for each music discovery approach for service developers to be able to further enhance music discovery experiences within these fields.

All but one of the proposed concepts work well for music discovery. The use of avatar characters and mood pictures for music discovery are the most encouraging based on the results. The results indicate that visual music discovery services can replace traditional music discovery services when designed to support different types of music discovery practices well.

The thesis contributes a generalized set of design implications for music discovery service developers, which has not previously been introduced in such magnitude for novel and playful music discovery services.

2. Related Work

This chapter presents related work for music discovery starting from the analysis of music files and retrieving the necessary information that can be used to build music discovery services. This is followed by a short introduction on what types of musical metadata is often used in user interfaces for music discovery systems. Music recommendation systems are typically an integral part of music discovery and the different recommendation mechnisms are briefly introduced here. The major part of the related work discusses the actual music discovery and categorizes a variety of applications in that domain to build a solid knowledge of existing state-of-the-art applications and services. Finally, user experience in music discovery is addressed.

2.1 Retrieving musical information and metadata

Music Information Retrieval (MIR) is an interdisciplinary research field targeting to broaden the understanding and usage of musical data through research, applications and tools using knowledge from computer science, music, signal processing and congnition (Casey et al. 2008). Schedl (2008, p. 17) describes MIR as follows: "*MIR is concerned with the extraction, analysis, and usage of information about any kind of music entity (for example, a song or a music artist) on any representation level (for example, audio signal, symbolic MIDI representation of a piece of music, or name of a music artist)*." A wide range of research communities from computer science, information retrieval, audio engineering, signal processing to cognitive science, psychology, and philosophy converge on MIR. Human-computer interaction (HCI) and user interface (UI) development are part of MIR topics (Casey et al. 2008). The challenges in MIR stem from music information being multifaceted, multimodal, multirepresentational, multiexperimential, and multicultural. The complexity of music information requires multidisciplinary research for reaching a robust MIR system (Downie 2003).

This section focuses on the content-based music information retrieval used in **P1** and **P3-P7** for music recommendations. Content-based MIR focuses on intelligent and automatic processing of music to try to help users in finding relevant music easily. An operational flow in content-based MIR firstly consists of a query from a user, then finding a match or a set of matches for the query, and finally formulating the output. Common MIR use case examples include music identification, detecting plagiarism and monitoring copyrights, identifying different versions of songs, finding similar sounding songs to a seed song, finding music belonging to a certain genre or mood, and finding and discovering music utilizing music recommendation algorithms. Three strategies are presented for solving the variety of use cases: 1) low-level audio features, 2) high-level music content description, and 3) metadata. The three strategies can support each other. For example, the musical key of a song or included instruments (high-level music content descriptors) can be stored as textual metadata for the song (Casey et al. 2008).

Low-level audio features are information retrieved from the actual audio signal. These low-level features are based on the analysis of short time segments of an audio signal, commonly segmented in three different ways: short fixed frames, beat-synchronous segments, and statistical measures. Processing of these audio segments is mostly done either in the temporal-domain (the natural domain of the audio signal) or in the frequency-domain. To better understand the basic audio signal operations, examples of low-level audio features in frequency and temporal domains are introduced with greatly simplified use cases in Table 1. The examples are based on information found from Casey et al. (2008), Mitrović et al. (2010), Peeters (2004), and Eronen (2001) and they include a broad range of audio features with detailed descriptions.

Low-level audio feature	**Domain**	**Description**	**Simplified use case example**
Pitch-Class Profile (Chromagram)	Frequency	Dividing the audio spectrum into frequency bands representing musical notes (more specifically chromas)	Finding tonality from the audio spectrum regardless of the octave
Spectral Centroid	Frequency	Finding a frequency that has an equal amount of energy below and above in the spectrum	Overall estimation of the audio signal having an emphasis on low, mid or high frequencies, i.e., brightness or dullness
Onset Detection	Temporal	Finding the beginning of the notes in a song	Can be used for estimating the activity and tempo of a song, and for automatic transcription of the music

Table 1 Examples of low-level audio features in frequency and temporal domain.

High-level audio features are typically based on analyzing the low-level features of the audio signal. They can be a combination of multiple low-level features used in a specific way. High-level audio features describe music with attributes that an experienced music listener would be able to identify from a song. High-level audio features include, e.g., timbre, melody, rhythm, pitch, harmony, key, structure of a song, and lyrics (Casey et al. 2008). Examples include finding verses and choruses from songs, defining the musical key of the song, and identifying the melody of a song. Calculation of these audio features from a large number of songs is time consuming and therefore often not done in real time. Luckily, these high-level features are constant and need to be calculated only once from each song. Therefore, one or more of these features can be written as metadata to complement the song and enable the use of this information in various services. The next section discusses musical metadata in more detail.

2.1.1 Musical metadata

Metadata is data that provides information about other data (Lehikoinen et al. 2007). Textual metadata embedded into digital music files is the most common way to access, search, and identify music collections (Casey et al. 2008). Basic musical metadata typically consist of pairs of descriptive attributes (e.g., genre, release year, artist name) and values for those attributes. Metadata forms the foundation for many music discovery applications regarding the availability of musical data to be used for searches and the user interface.

The most commonly used metadata format for music files is ID3 (1998). An ID3 tag can be embedded as a part of the actual media file to describe its contents. The current ID3v2 is supported by the majority of music players and services. ID3 tags were originally developed for MP3 files but are compatible with various other formats. Other music and multimedia metadata sources include, e.g.:

- MusicBrainz (MusicBrainz 1999): Online moderated database for metadata with track level identification.
- Gracenote (Gracenote 2000): The largest online service for music and video metadata
- Discogs (Discogs 2000): Online database of electronic music releases and related metadata

For a more detailed description and comparison of several metadata sources, see Corthaut et al. (2008).

Metadata for individual songs can be defined using content-based audio analysis (see previous section and Dunker et al. 2008), human annotation, or hybrid approaches combining one or more approaches (e.g., Hu et al. 2009). Metadata generated using content-based audio analysis is automatic and fast compared to human annotation and can be used to analyze music catalogs of millions of songs. Due to the complexity of polyphonic audio signal and analysis algorithms, the results are prone to errors. For example, it is extremely difficult to transcribe a song or detect the instruments, mood of a song, or the gender of the singer accurately (e.g., Hamel & Eck 2010, Ellis et al. 2013, Kim et al. 2010, Turnbull et al. 2008). The Music Information Retrieval Evaluation eXchange (MIREX 2014) performs a yearly collective evaluation of existing algorithms and is a good source for state-of-the-art status for related algorithms.

Human-annotated metadata can be inserted by end users, content creators, or companies. Due to the nature of human-annoted data, the metadata can be either factual or influenced by subjective opinions, culture, or false assessment due to limited knowledge. For example, artist or song name data are factual forms of metadata having an explicit original representation set by the artist or their label. Still, it is common that in human-annotated songs the names are misspelled by accident or the content is formulated differently than, e.g., in some commercial databases. Genre information is an example of a subjective and, cultural metadata attribute and can vary greatly throughout the world. Examples include the definition of world music in Finland and in India, and the categorization of a song as pop or rock in borderline cases. Currently, a common way of adding metadata to multimedia content is folksonomies: to add collaboratively generated, user definable labels describing the content (Lehikoinen et al. 2007), commonly known as social tags. Social tags can be gathered from services such as Last.FM (2002) or harvested automatically from relevant web pages.

Many companies are making a business from accurately hand-annotated metadata, using it for various services or offering it to others. One example from the music discovery domain is Pandora (2000), which offers personalized music recommendations in the form of a radio stream using a rich and extensive metadata library of over 800 000 songs. In Casey et al. (2008), it was estimated that entering such metadata manually for one song takes 20-30 minutes from an expert annotator. Another relevant source for expert annotations is AllMusic (1991), specializing to genre and mood information.

Inconsistencies and errors in metadata can produce problems in music applications and services. Inaccurate metadata can lead to problems, for example in finding the songs with appropriate queries (e.g., artist name, track name, album name), providing accurate recommendations that use metadata, and user interface problems when presenting the metadata.

2.2 Using musical metadata

The previous section described ways to retrieve and generate metadata for music files. This section presents an overview of the most common types of metadata used in music discovery services and examples of how this metadata is presented to the user. Visualizing musical metadata within the user interface is one of the key elements when designing for a good and entertaining user experience for music discovery services. For general information regarding information visualization, see Holm (2012, Chapter 3). Musical metadata is also a fundamental block for music recommendations. In addition to the use of musical metadata, other metrics can be used to aid visualization and recommendations. For example, popularity, the amount of followers or fans, can easily be combined with musical metadata such as artist information, which would be difficult to visualize as such.

2.2.1 Genre

Musical genres are commonly used for categorizing music. Depending on the genre classification source used the categorization can be rather high-level such as rock or more fine-grain, e.g., symphonic rock. Musical genres are labels describing music with no exact definitions and boundaries. Musical genres have been developing as a product of interaction between public, marketing, historical, and cultural factors. Music within a genre often shares certain common aspects relating to, e.g., rhythmic structure or instrumentation (Tzanetakis & Cook 2002). Musical experts typically create the overall structure for musical genre categorization. The structure can be populated with music either manually or with automatic genre classification. There are many different sources of genre structures and they are not consistent. Examples for genre listings include AllMusic (1991) with 531 genres and ID3 (1998) with 226 genres including extensions.

Taking into account the huge number of musical genres, manual music classification is a difficult and time-consuming task. Due to many factors, e.g., subjective and cultural factors to name only a few, the classifications are often not consistent. Commercially produced and distributed music arrives to online music services with genre labels attached by the record companies and the variation in genre classification can be very different between companies.

The common aspects in music within a specific genre are a starting point for automatic genre classification. A great number of research contributions are available relating to automatic genre classification. Overall, automatic genre classification with small datasets and higher-level genres can achieve approximately 73% accuracy (varying from 55-90% depending on the genre) (MIREX 2013). A good overall discussion regarding automatic genre classification can be found in McKay & Fujinaga (2006) and a more critical view by Sturm (2012).

The inconsistency in music genre classification and the great number of genre alternatives pose a potential problem for the end user with music discovery services. An ordinary user without a specific musical knowledge is often unable to comprehend and recognize a great number of different genres and sub-genres. Despite the problems with using musical genres for music categorization, they are the most common way of categorizing musical content in music discovery services.

Musical genre can be visualized in many different ways. Examples include textual descriptions of the genre, colors, icons and textures, fonts, avatars, and virtual world appearances. A more detailed description and comparison of different genre visualizations can be found in Holm (2012).

2.2.2 Tempo

The tempo of a song can be described by an analogy to the listener's "*foot-tapping or hand-clapping with music, tasks of which average human listeners are capable*" (Dixon 2001, p. 1). The tempo in a song is indicated by beats per minute (BPM). The tempo of a song can be automatically calculated (e.g., Dixon & Cambouropoulos 2000, Alonso et al. 2004, Eronen 2009) and the value can be retrieved from services providing metadata for songs, e.g., Gracenote (2000) or The Echo Nest (2005).

The exact value of a tempo in a song can be used to sort songs based of their speed. Commonly, the exact value is not exposed to the end user but categorized, e.g., slow, medium, fast. Other alternatives include offering a slider or other user interface controls to adjust the suitable tempo or a tempo range, and returning the matching songs. Based on the use case, a set of songs are commonly displayed or filtered away in the user interface. Difficulties may arise with songs that do not have a constant tempo (or no pulsation at all). Alternatively, the textual values for the tempo can be replaced with icons or pictures. From the research side, there are some two dimentional x-y space mappings of music where tempo is used as one attribute on one axis (Van Gulik & Vignoli 2005). In Holm (2012, Publication 1), a broad range of different graphical objects (geometrical shapes with effects and colors) are associated with musical tempo and evaluated by users.

2.2.3 Release year

In many music discovery systems the release year of an album or a song is used as an attribute for searching. Also, in many cases the release year can be shown while viewing detailed information regarding a song. The results of searches made using the release year are often clustered to decades and centuries or eras (classical music). The decades are often used in combination with genres e.g, “80s rock music”. A more academic and experimental approach for visualizing the release year is found in Holm (2012, Publication 2). A broad range of different graphical objects (geometrical shapes with effects and colors) are associated with the release year and the association is evaluated by users.

2.2.4 Mood

Music listening may evoke strong emotions in people and can have an effect on people’s mood, facial expressions and physiological reactions (Juslin & Sloboda 2001). As a consequence of its effectiveness, music is used for mood enhancement in many applications including sports, TV, relaxation, movies, and therapy. Virtually every brain region mapped by cognitive neuroscientists is activated by music. Music has access to regions involved in, e.g., planning, motivation, forming expectations, memory, association, and attentional systems (Levitin 2007). In addition to mental activity, music can have other physiological effects on a person including sweating, respiration and heart rate (Levitin 2007). Most people are able to select the right type of music for their emotional goals, e.g., “*people in a state of unpleasantly high arousal (for example, while driving in heavy traffic) generally prefer quiet, relaxing music, while people who are in a state of pleasantly high arousal (for example, exercising, working out) will prefer loud, energizing music*” (Levitin 2007, p. 3). As music can play such a significant role in evoking, expressing, and communicating emotions in both listeners and performers (Juslin & Sloboda 2001), it is natural for people to also categorize music based on emotions and/or moods.

Sloboda & Juslin (2001) describe the main differences between emotions and moods. Emotions last for a short while, have an identifiable stimulus, and are complemented with distinct facial expressions. Moods may be experienced for a long period of time and cannot be directly expressed by a person. Any specific object or situation does not create a mood in a person. These terms are commonly mixed in MIR literature.

In MIR systems, emotions are commonly classified using categorical or dimensional approaches, which have roots in psychological research. It remains a challenging problem to find a perfect set of emotion labels to be used as a base set for applications and algorithms (Kim et al. 2010). A study by Skowronek et al. (2006) pursues a base data set by evaluating 470 music clips from 12 genres with nine bipolar mood scales. The results show only half of the clips receiving strong judgements for the moods meaning that only half of the clips were easy to judge. Sloboda & Juslin (2001, p.76) explain the concept of basic emotions as; “*there is a limited number of innate and universal emotion categories from which all other emotional states can be derived*”. There is a big variation in different emotional categories depending on the use case and application area. Ekman (2004) proposed six emotions including anger, disgust, fear, joy/happiness, sadness, and surprise. These basic categories are universally recognized. The number of emotional categories tends to be higher in automatic mood classification than in most psychological research (Ekman 2004), although there are exceptions such as a study by Zentner et al. (2008) with a significant amount of mood classes.

Social tags of songs are commonly interpreted to determine different emotional categories as presented in, e.g., Hu et al. (2007). The later work by Hu et al. (2009) extends the amount of mood categories to 18, containing, in total, 135 mood tags. The All Music Guide (1991) features 179 distinct mood labels. However, the labels are partially related. In most of the music discovery services that are targeted for end-users, the amount of mood labels is typically smaller.

An alternative way of modeling emotions focuses on "*identifying emotions based on their placements on a small number of dimensions, such as valence, activity, and potency*" (Sloboda & Juslin 2001, p. 77). The most commonly used dimensional scales for emotion are by Russell (Russell 1980) and Thayer (Thayer 1989). The Russel's circumplex model (Figure 2) maps the x-axis to valence and the y-axis to activation. The model is built in such a way that opposite emotions face each other.

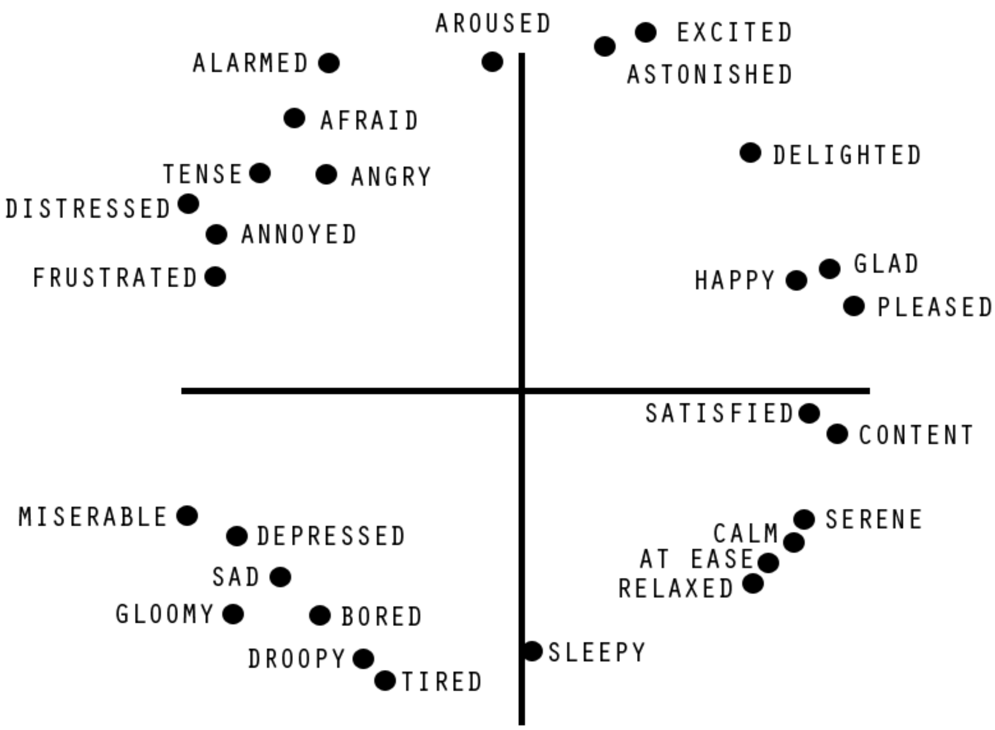

Figure 2 Russell's circumplex model of emotions. Adapted from Russell (1980).

In MIR, the most typical representation of musical affect is Thayer's two-dimensional valence-arousal space (Thayer 1989), which is derived from Russell's general model of human emotions (Russell 1980). Thayer's model represents emotions in music using two dimentions: valence and arousal. The valence dimension describes positive and negative emotions, whereas the arousal dimension represents the amount of energy from low to high. Based on the model, music can be roughly categorized into one of these four quadrants: low valence and high arousal (anger), low valence and low arousal (depression), high valence and high arousal (joy, exuberance), high valence and low arousal (contentment). Kim et al. (2010) present a state-of-the-art review regarding music emotion recognition for further reading.

Moods are typically used in music discovery applications as textual labels, descriptive pictures or icons. Moods enable the users to categorize music beyond genre boundaries and receive a selection of potentially very different types of songs matching the selected mood. Further filtering can be done using other types of metadata, e.g., genre or era to achieve more consistent playlists for the user. See more details and examples in Section 2.4.9.

2.3 Music recommendation systems

Music recommendation systems play an important role in music discovery. Music recommendation systems work in the back-end of music discovery services and are able to offer new music for users based on a query, music consumption history, and their preferences. In Celma (2008), music recommendation methods for individual users are divided into five categories, which are summarized in this section: demographic filtering, collaborative filtering, content-based filtering, context-based filtering, and hybrid methods. For a slightly different categorization of music recommendation methods and a comparison, see Burke (2002).

2.3.1 Demographic filtering

Demographic filtering aims to cluster groups of users that like a certain song. This clustering is based on first trying to identify a representative and stereotypical person that listens to a specific song by analyzing personal data, geographic data, and psychographic data. The analysis data may include, e.g., age, gender, country, and interests. The method can be described by: “X and Y are similar in age, gender, country and their interests match”. The system then recommends similar songs to people having a similar demographic profile. In some cases these types of recommendations can seem too general and not personalized enough. Additionally, this method requires the user profile to be present in order to produce recommendations. There are attempts to automate user profile creation by harvesting information from a user’s Internet content (Pazzani 1999, Celma 2008).

2.3.2 Collaborative filtering

Collaborative filtering systems in music predict affection towards music by connecting users with a similar listening history, preferences, or feedback, and then recommending songs between like-minded persons. (Herlocker et al. 2000) The method can easily be simplified as “people who listen to X also listened to Y and Z”. The filtering is based on human actions instead of content-based analysis of music. The benefits of collaborative filtering include being able to filter any types of content (e.g., music, books, movies, art). The methods can also be used between different types of content and the method is not prone to data analysis and retrieval errors as content-based methods are. (Herlocker et al. 2000) The potential problems with collaborative filtering include a “cold start” problem where no recommendations can be made between the users if a) there are not enough users, b) there is not enough similar data between the users. Collaborative filtering has a strong popularity bias and is not able to recommend new and unheard music for users or music from the long tail (Celma 2008).

2.3.3 Content-based filtering

Content-based filtering commonly refers to analyzing the audio signal for low-level and/or high-level features (see Section 2.1 for details) and using these attributes to describe the content. These features, such as timbre, tempo, key, and rhythm can be used for recommending a set of similar type of songs from a music catalog. Content-based filtering provides information such as: “X and Y sound similar”. Most of the research work in MIR focuses on content-based analysis of music (Orio 2006). Different audio features can be used separately or as a combination to match the user preferences. A common use case for content-based recommendations is to give a seed song and retrieve a playlist of similar sounding or otherwise similar songs. In this use case, content-based filtering does not have the cold-start problem and thus does not require many users or user profiles to work. In Orio (2006), the basic assumption for using content-based filtering is that the metadata for a song is unsuitable, unreliable, or missing. Content-based filtering can be, and often is used to complement other music recommendation methods (see Section 2.3.5) and can bring new discovery experiences from the long tail (see Section 2.4).

The disadvantages with content-based filtering include, e.g., the time-consuming analysis of musical content, the unreliability of the analysed features due to the domain complexity, and the ability to produce only similar item recommendations not taking into account, e.g., popularity or community preferences. Celma (2008) broadens the definition of content-based filtering to include processing manual annotations by domain experts or using tags from community users describing the actual content.

2.3.4 Context-based filtering

Context-based filtering focuses on cultural information for computing artist or song similarity. The techniques involved include web mining and social tagging data for finding similar items or users (Celma 2008).

Web mining can be divided into content, structure and usage mining. Web content mining refers to analyzing the actual web page content. Web structure mining follows links on web pages, i.e., network topology analysis. Web usage mining uses session logs to retrieve information regarding, e.g., users' habits and preferences (Celma 2008).

Social tagging is a practice of attaching keywords to classify and describe content. These keywords can be freely chosen. Gathering and analyzing these tags with different methods is a common procedure for context-based filtering.

Context-based filtering is able to take into account the cultural aspects in music classification and web crawling enables retrieval of what is trendy and popular at a given time. The limitations include: a potentially narrow scope of data focusing on popular items, the common tags used for many types of content, and the constraints of a controlled vocabulary (Celma 2008).

2.3.5 Hybrid methods

Hybrid methods combine one or more of the previous methods for gaining a better overall performance. Different ways of combining the recommendation algorithms are described in Burke (2002). These include:

- **Weighted**: combine a weighted output from different recommendation methods
- **Switching**: use a criteria to switch between different recommendation methods
- **Mixed**: expand the dataset using input from different recommendation methods and recommend based on the larger dataset
- **Cascade**: applying one recommendation method first to rank the items and using another method to re-rank or refine the results

2.4 Music discovery and related applications

An important use case for MIR technologies is music discovery, i.e., finding appropriate and/or new music for the listener. In Cunningham et al. (2007) everyday encounters with new songs was explored using a diary study with 41 students from a University HCI course. The users marked down all incidents that brought them in contact with new music during three days. Each user had, on average, 3.3 new music encounters each day with over 62% of them being perceived positive. Roughly one third of the encounters were active, i.e., where the users were specifically seeking for new music. The dominant media for music discovery in the study were the Internet, radio and TV. The participants encountered new music at many locations and at all hours of the day and night. The study highlights the big amount of daily music encounters and the potential of music discovery applications for better discovery experiences compared to the traditional media.

The Long Tail in music is relevant for music discovery and is described by Celma (2008). The Long Tail consists of the head including popular, trending, and frequently listened to "hit" songs and the tail featuring a large number of unpopular and many times never accessed songs (Figure 3). In 2007, only 1% of all songs accounted for 80% of all music sales and 1000 albums accounted for 50% of all album sales.

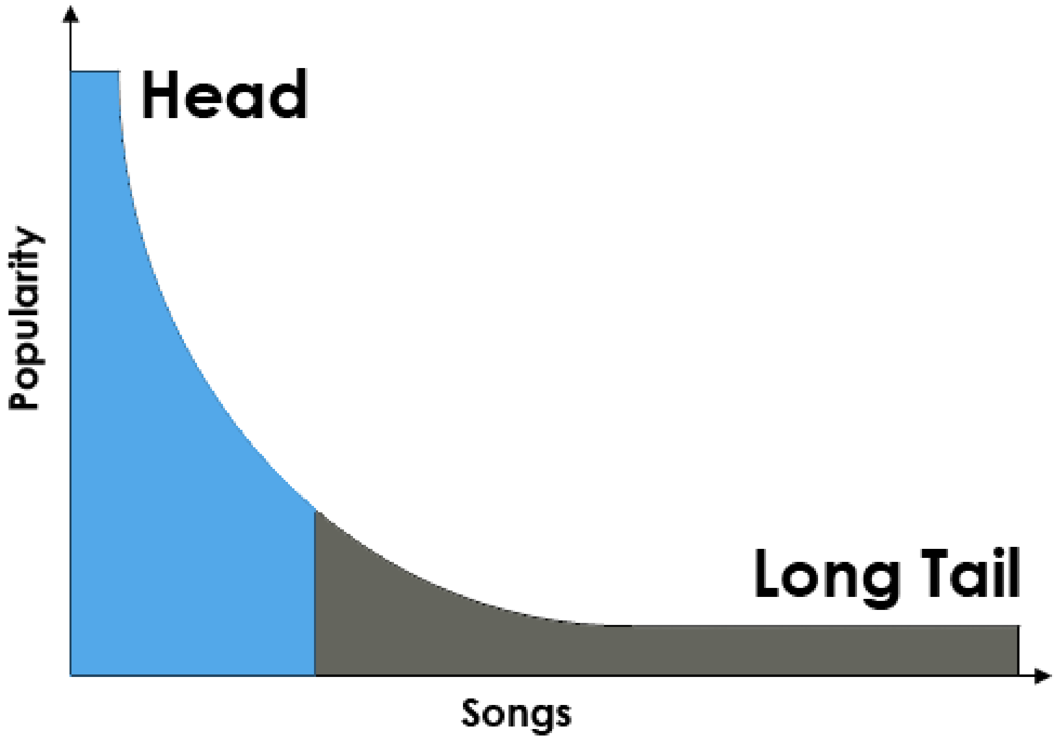

Figure 3 Illustration of the Long Tail in music, where the majority of music is undiscovered by the masses.

It is quite certain that the "heavy tail" (Celma 2008, p. 79) includes good and relevant music for consumers but work is needed with music discovery services to find good and entertaining ways of reaching the whole spectrum of music. Celma & Herrera (2008) propose the use of content-based retrieval for achieving this. Spotify (2006) is one of the most popular music streaming services at the moment featuring 4 million songs that have never been played. Now there is a separate web service to access those forgotten songs (Forgotify 2013).

Lillie (2008) describes three different types of music discovery: **active, passive, and exploratory**. Active music discovery is about searching through music catalogs. The user has at least a coarse idea what to search for. In the case of music this can mean doing searches with the artist name or partial lyrics. These searches are often reflected in the metadata of the songs, i.e., anything available in the metadata can be searched easily. Passive music discovery is about music recommendations, particularly the use of different music recommendation systems. The music recommendation systems may work automatically using a user's earlier preferences and music consumption, context, activity, or social connections. In other cases the recommendation system takes input from a user, e.g., a seed song or mood to generate a list of recommendations. Exploratory music discovery can be motivated by enthusiasm and curiosity instead of a need for some exact song. Exploratory music discovery often involves browsing music catalogs with various means without precise expectations of what will be found.

In a survey regarding the everyday-life music-seeking behavior of 15 young adults (Laplante & Downie 2006) the role of non-goal oriented browsing of music and serendipitous music discoveries were emphasized. The results suggest putting more focus into music discovery and novelty-biased MIR systems in the future. Brown et al. (2001) make an analogy between music discovery and the physical shopping experience of buying new music. Currently, there are many 'shops' to choose music from, i.e., hundreds or even thousands of services available supporting music discovery in different ways. Several examples are categorized and introduced in the following sections to give an overall picture of state-of-the-art music discovery applications and services.

2.4.1 Versatile music discovery services

Versatile music discovery services combine a wide set of features for music discovery and often include 3rd party application support to further extend the discovery opportunities within the service.

Spotify (2006) is one of the leading online streaming music services, featuring over 20 million songs, music searching, browsing music based on different categories (e.g., moods, top lists, new releases, genres, eras), discovering new music based on music consumption history, tailored radio stations, and support for third party applications inside the Spotify application utilizing the massive music catalog (Figure 4). Third-party applications offer many new opportunities for music discovery and there are over 100 alternatives available as of this thesis' publication. Many popular music discovery services have created their own Spotify applications, e.g., Moodagent (2001) and Last.fm (2002).

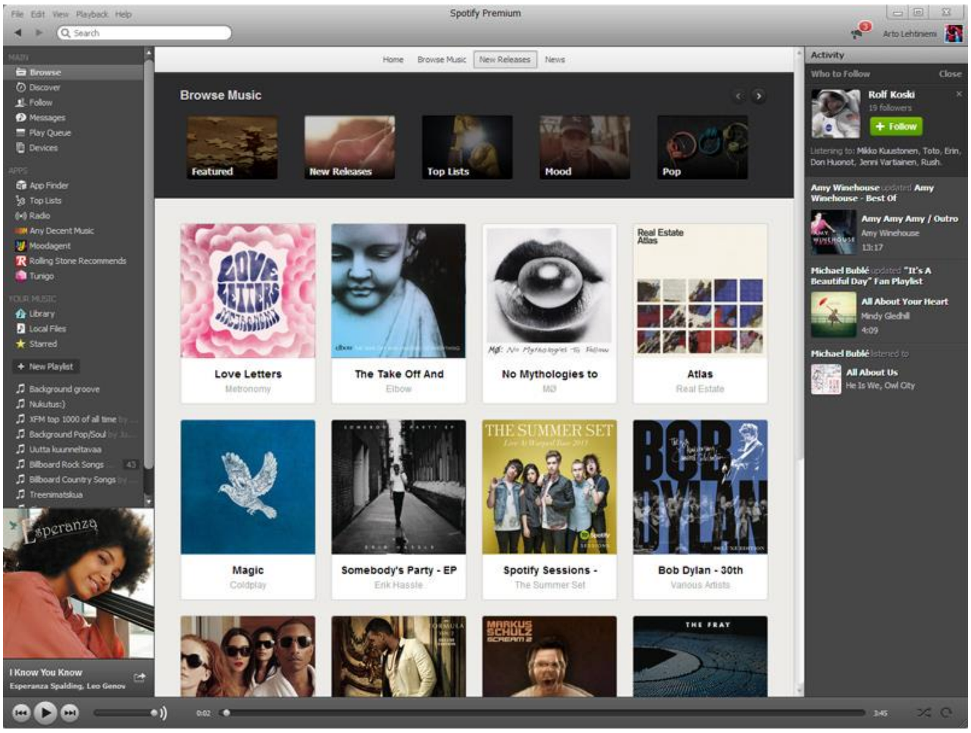

Figure 4 Spotify (2006): Versatile music discovery system with 3rd party application support.

Deezer (2007) is another versatile music streaming service with charts, radio channels, and similar artist recommendations utilizing a 30 million-song catalogue. Like Spotify, Deezer also includes a wide range of 3rd party applications for music discovery.

Grooveshark (2007) is a popular online music streaming service with streaming radio and music recommendations in addition to search for music, friends, and user-initiated live broadcasting of music. Other users are able to follow these user-initiated broadcasts as if listening to a radio station and chat with other listeners. One special feature about Grooveshark is that all of the music is coming from their users. The system analyzes users' music libraries and adds them to their online library to be used in the service.

2.4.2 Radio-based music discovery

Radio-based music discovery refers to solutions where the music is automatically chosen or a playlist/stream is automatically generated based on minimal user input. In these solutions the user is able to listen to the music but does not have full control over the playlist i.e. is not able to choose songs and artists freely or is only able to skip a certain number of songs.

Pandora (2000) is one of the most popular music recommendation and discovery services available in the United States, Australia, and New Zealand. The service has a large catalog of hand-annotated music with over 400 different musical attributes. The user is able to use a song or an artist as a seed to create a new channel and the system will play recommended songs based on the seed. The user is able to vote thumbs up or down to audible tracks and thus change the course of upcoming music. The detailed, state-of-the-art handmade music annotation helps Pandora to be highly accurate with their music recommendations.

Nokia MixRadio (2011) is included with recent Nokia mobile devices and offers over 300 ready-made playlists covering a broad range of different musical styles using a database of over 20 million songs (Figure 5). The user is able to skip a certain amount of tracks but not able to search and play individual tracks à la carte. Nokia MixRadio also incorporates information about upcoming concerts and enables users to purchase tickets through the application and get navigation guidance to the concert venue.

8tracks radio (2008) is an internet radio service featuring user-curated playlists. The user is able to choose any mood, genre, or activity to start selecting a playlist. The system then lets the user further define their search with additional words resulting in a playlist to listen to.

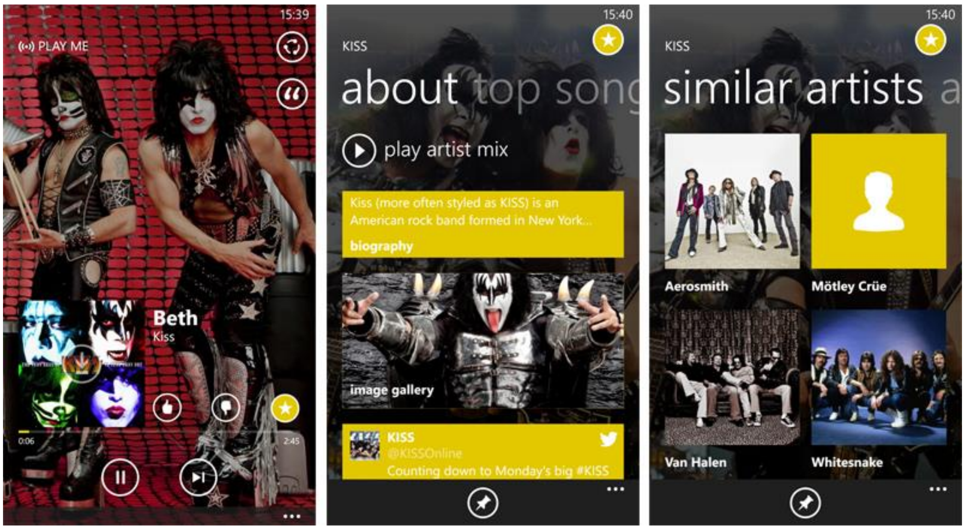

Figure 5 Nokia MixRadio (2011): Radio-based music discovery with the ability to choose a style of music but not the exact songs (or artists) to listen to.

2.4.3 Handpicked music recommendations for music discovery

Handpicked music recommendations often feature expert picks from music publications and blogs or radio stations. The reviews and descriptions are typically relatively detailed and thus the number of such recommendations is commonly small. This method of music discovery is very similar to recommendations from friends and they often capture a user's attention better than a mass of automatic recommendations.

KCRW Music Mine (2013) is an iOS and Spotify application provided by the KCRW radio station. The application features daily changing handpicked songs to listen to while reading the artist background information. The radio station is known for introducing new (soon to become famous) artists to listeners. The application additionally includes selected music videos and music recommendations based on the songs listened to (Figure 6).

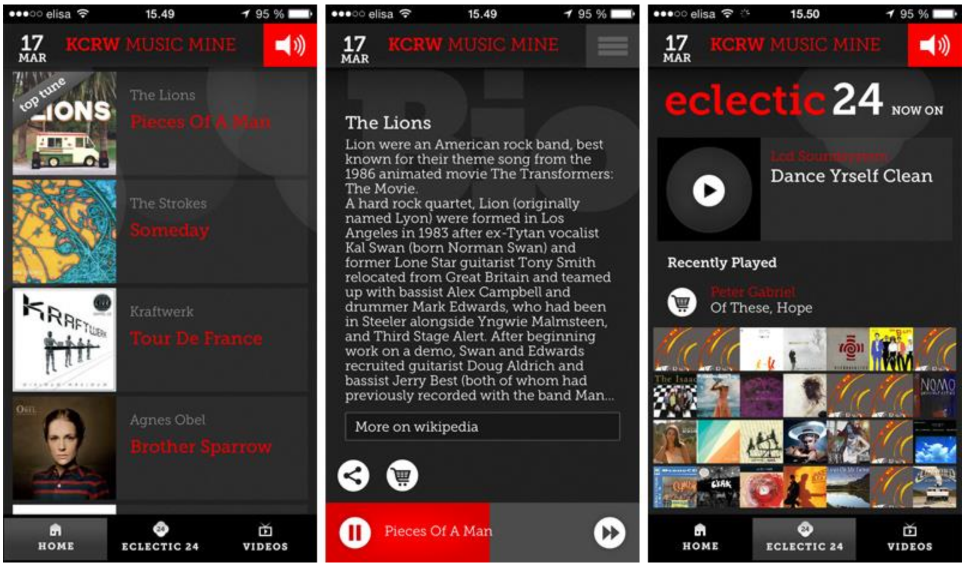

Figure 6 KCRW Music Mine (2013): Handpicked music recommendations for music discovery including detailed reviews and related information.

Stereogum (2002) is a famous MP3 blog, an online outlet for indie and alternative music typically featuring unknown artists who later become famous. The service includes reviews, news and stories with linked music for the discovery of new artists.

Beats music (2014) could easily be categorized as a versatile music discovery service due to its monthly subscription to stream millions of songs, social features, and many ways of discovering music. Still, the key for music discovery in Beats music is the emphasis on its suggestions, highlights and curated playlists. Beats music is collaborating with several famous music magazines and web sites (e.g., Academy of Country Music, Friends of Beats, Pitchfork, Rap Radar, Rolling Stone) to acquire curated data to be used in the service. The picks from each curator can be accessed separately. A new music discovery feature was recently added to Beats music allowing users to modify sentences for music discovery. The Sentence feature enables the user to choose where they are, what they feel like doing, who they are with, and a type of preferred music. A playlist is automatically generated based on the selected attributes.

2.4.4 Community-based music discovery

Community-based music discovery is all about social music services. The music related data is processed over a large number of people (or related sites) and compiled as alternatives to new music discovery.

SoundCloud (2007) is an online music and audio distribution platform originally designed for grassroot artists to upload their music to the Internet using SoundCloud widgets (Figure 7). The service has grown fast and now features a large community of artists and ways to discover new and trending music. The music in the service is displayed as waveforms and one special feature includes the ability for users to leave timed comments to the songs regarding a specific moment (see the waveforms in Figure 7). These comments are displayed while playing the song and bring a strong community sense for new music exploration.

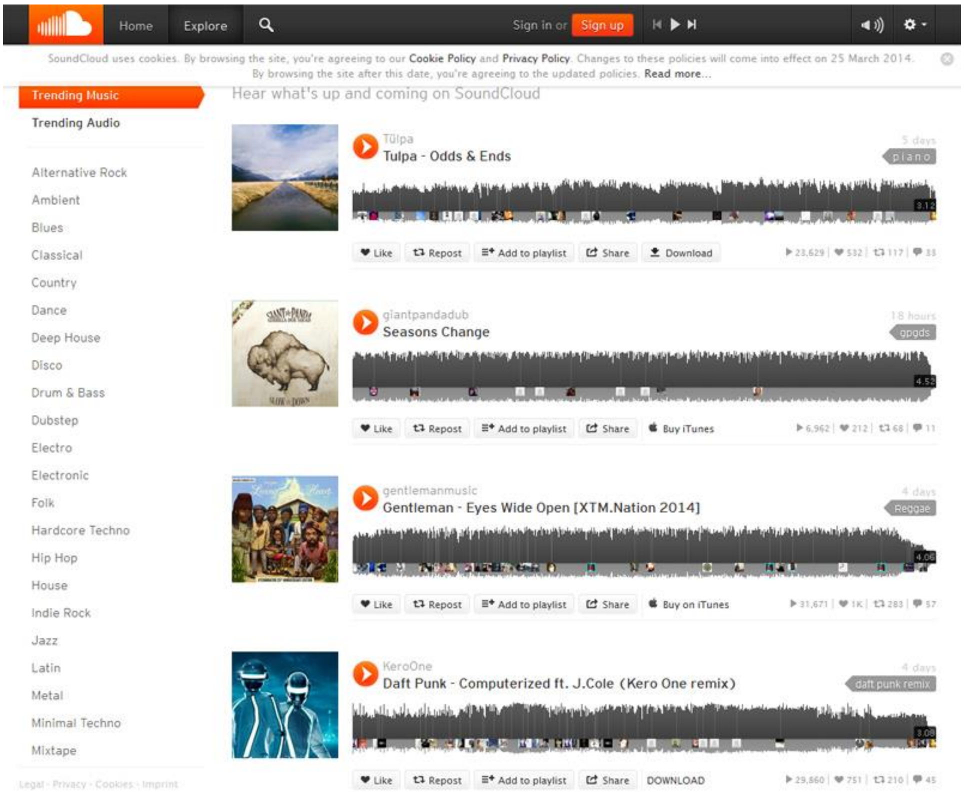

Figure 7 SoundCloud (2007): Community-based music discovery with the ability to upload music from artists and receive user comments regarding individual parts of the song.

The Hype Machine (2005) combines information about music discussed in blogs. The harvested information is categorized according to latest, popular, different genres, specific blogs, and exclusive premieres. The service enables users to listen to the tracks and read associated blog postings. Basic social features are also included featuring liking and posting to common social networking services.

Last.fm (2002) has taken a slightly different approach to community-based music discovery with large-scale collaborative filtering. Last.fm logs a detailed user profile of consumed music in various different music services and applications using an

AudioScrobbler. Most popular services and music player software support “scrobbling” track information to Last.fm and this data is used to recommend new and relevant music that other community members have listened to. iTunes Genius (2008) is another example of using collaborative filtering based on a listened track. iTunes Genius takes in a seed song and automatically creates a playlist from a user’s local tracks that best match the seed song.

Rdio (2010) is a socially driven music service which seamlessly links to Facebook and Twitter. With Rdio the user is able to search and listen to songs, explore music recommendations based on a user’s activity and recent listening, browse through community favorites and most played tracks. Users are able to follow social networking friends and their music listening and discover new music.

This is my jam (2013) lets users create their “jams” out of one (and only one) special song as follows. The system lets the user search for the song, choose a suitable visual appearance, and attach a note describing why it is special. The user is able to follow other people and friends in the service and listen to their creations, i.e., jams. The system will also automatically recommend jams featuring the song used in their own jam.

Shuffler.fm (2010) lets users discover new music and related news shared in blogs in a visual way using large graphical elements from album covers, music videos, and web site logos. The service integrates a number of music blogs and sites for music discovery in addition to browsing music from a specific style or based on popularity. Shuffler.fm also features a radar that can alert a user when new music is available from selected artists, sites, or people.

2.4.5 Contextual music discovery

There are many definitions for context, ranging from synonyms, e.g., environment and situation (Salber et al. 1999) to more detailed descriptions. Context is defined for application and service scenarios in Dey (2001, p. 4) as: “*Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.*” In mobile music services, location and time are the most commonly used attributes of context. The use of activity detection for context recognition in various applications has gained attention during recent years.

Soundwave (2011) keeps track of music listened to with a mobile device’s native player or several big streaming services. Music discovery is the main focus of Soundwave (Figure 8). Users are able to follow one another’s activity feed as they listen to music and view that activity on a map interface. Soundwave users are visible on a map and the application enables users to circulate an area of the map to see what has been listened to in that particular area.

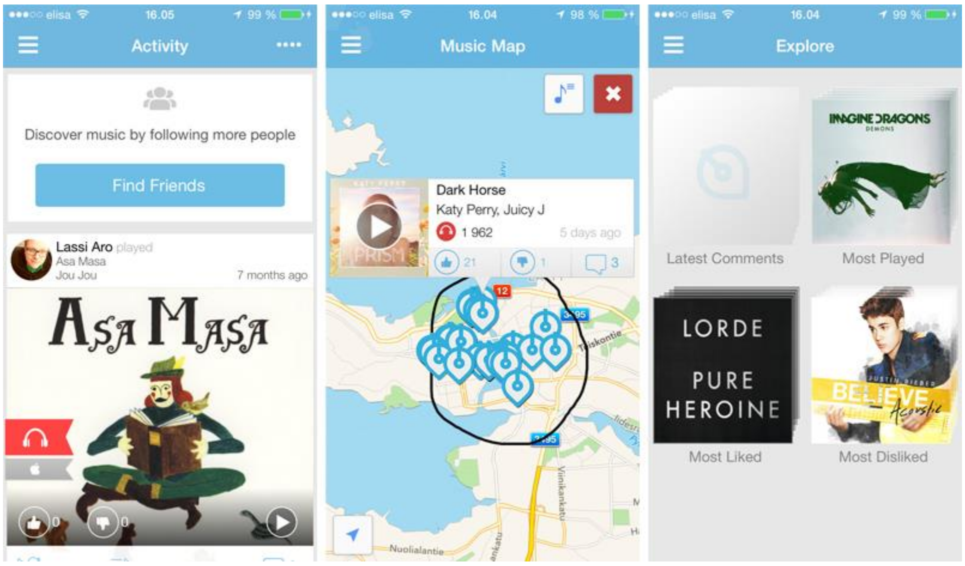

Figure 8 SoundWave (2011): Contextual music discovery with the ability to select music from a map interface (middle) and explore music contextually (right).

ContextPlayer (Okada et al. 2011) is a context-aware music player for situational recommendations. The user is able to listen to music with the application and build a user profile. The system tries to adapt to the context of the user's musical preferences and recommend new songs. The user interface is mainly based on small album covers and colored tiles. A common cold-start problem (i.e., when the system does not know about a user's musical preferences yet) is handled by initially providing recommendations from top selling songs in the corresponding area. The ContextPlayer was evaluated by Okada et al. (2011), showing that the users had problems comprehending the situation's concept and its link to the recommended music. The majority of users enjoyed the application and described it as being "cool" and having "wow-factor". Many users indicated their need for a search feature and wanted to choose specific songs to play.

Foxtrot (Ankolekar & Sandholm 2011) addresses location-based music discovery in terms of creating a soundtrack for a user's location. The prototype system first retrieves songs, audio comments, and ambient sounds that have been listened to in the given location. Next, the system analyzes these audio files in terms of likelihood of being suitable for the user. Finally, the system re-arranges the audio files to achieve a pleasant-sounding mix to enhance the user experience. The output is a radio-like stream of audio while the user is on the move. The desktop interface uses Google Streetview (Google Streetview 2014) and the mobile interface uses Google Maps to show the music on a map. The related user evaluation indicates that music is the most engaging content for such prototype. Music resulted in emotionally richer reactions compared to ambient sounds in the evaluated prototype.

Ambient sounds were seen to provide a better sense of surroundings while virtually exploring the streets. Overall, the study suggests that a location-aware music system can provide engaging mobile experiences.

A contextual music recommendation prototype for daily activies and a related user study is presented in Wang et al. (2012). The developed system classifies sensor data on a mobile device to classify activities and recommends suitable music automatically for the user. The user evaluation was mainly concerned with the recommendation algorithm stating that the system was able to avoid cold-start problems in music recommendations using contextual information and readily annotated songs for the activity classes. The paper also states that although contextual annotations are subjective, many users agreed with the annotations in the prototype system. More background information related to context recognition using different sensors can be found from Leppänen & Eronen (2013), Eronen et al. (2006), Lester et al. (2006), and Ming et al. (2010).

2.4.6 Audio recognition-based discovery

Users often encounter new and interesting music in their daily lives that they do not recognize. Services in this category help the user to identify music with various metrics and after a successful recognition, they offer a variety of added functionality.

Shazam (1999) is a mobile service that captures an audio clip, calculates an audio fingerprint from the clip and compares it to fingerprints from a database in order to find a match and thus recognize the song (Figure 9).

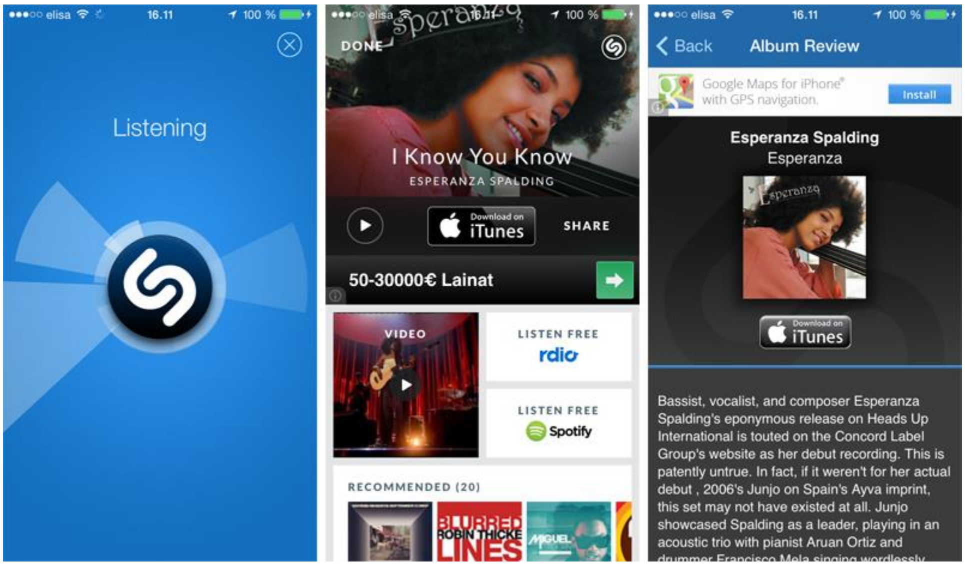

Figure 9 Shazam (1999): Music recognition-based discovery where a song is recognized using a microphone (left) and the results and related information are displayed (middle and right).

The recognized song can then be listened to using major services such as Spotify (2006) and the application can recommend similar songs to be listened to or nearby performances from the associated artist. The application keeps track of recognized songs and includes posting to social networking sites. Wang (2006) has published a detailed description on Shazam.

SoundHound (2009) extends the recognition of songs to playing, singing and humming. Like Shazam (1999), the application also links recognized songs to streams from major services and additionally provides links to lyrics and available videos. The application also provides community based charts for new music discovery.

2.4.7 Discovery by textual query

In music discovery by query the user either knows a good song, artist, or album to search for and the service provides a list of relevant recommendations based on the query. In some cases, the queries can be more complex and include filters, e.g., “90s pop music” and the recommendations are displayed accordingly.

Youtube (2005) is a well-known video service that also includes a large amount of music. The user is able to make queries and the system recommends additional related content on the sidebar (Figure 10).

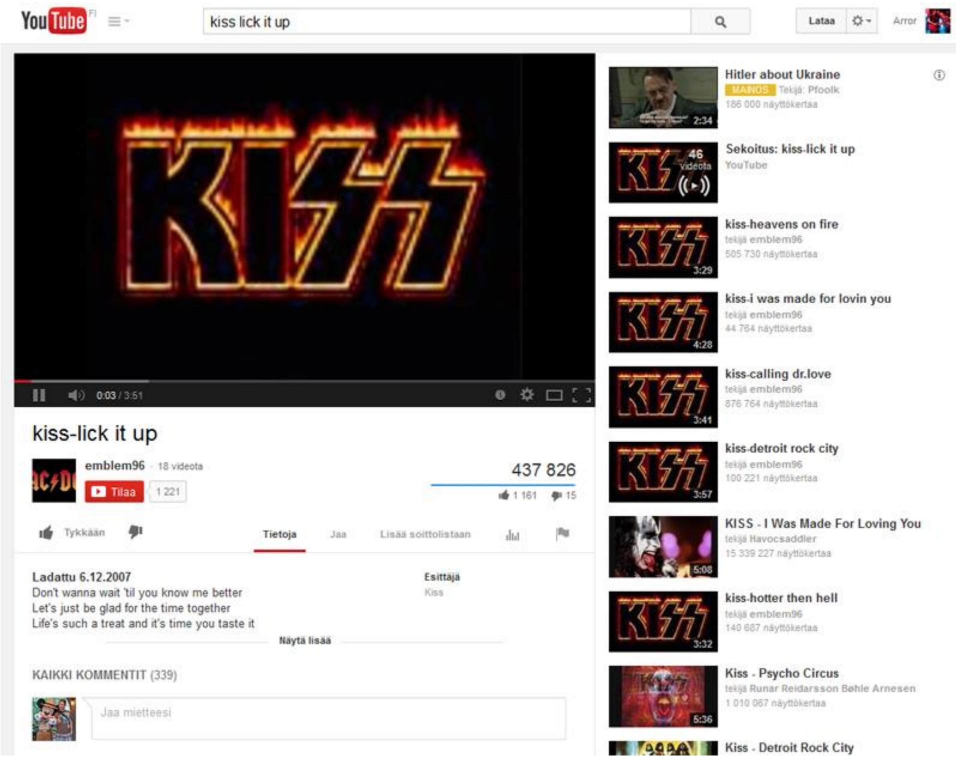

Figure 10 Youtube (2005): Discovery by textual query featuring related content in the sidebar.

There are many mashups using Youtube content. One good example for music discovery is Seevl (2012), which enables the user to search for music providing more related information, e.g., biography, top tracks, genre, members, and discography by harvesting common data sources such as Wikipedia (2001), Freebase (2007), and MusicBrainz (1999).

TasteKid (2009) provides a complete entertainment recommendation system where users can query music, movies, games, books, and shows and receive similar recommendations instantly.

2.4.8 Discovery using bubbles, radial and other geometrical shapes

This section presents selected applications and services that enable music discovery using different geometrical shapes. These applications have a strong and non-traditional user interface focus for music discovery and exploration.

Hitlantis (2009) is a service where the user is able to discover new music mainly from smaller and as-yet unfamous bands (Figure 11).

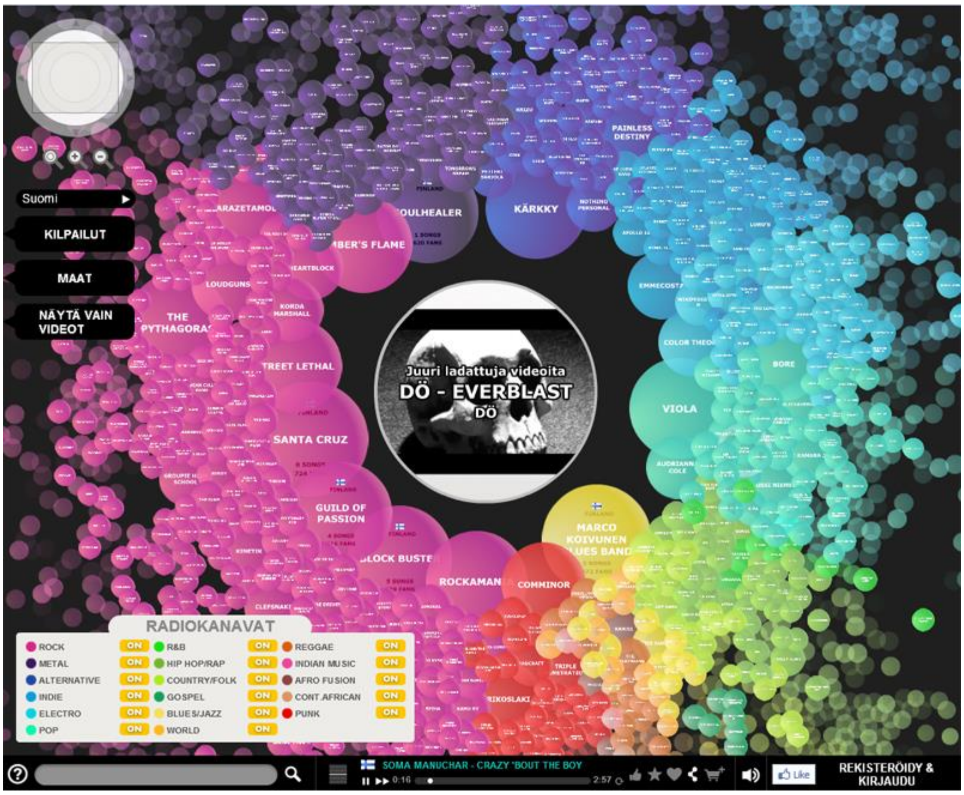

Figure 11 Hitlantis (2009): Discovery using bubbles representing artists with color-coded musical styles.

The bands, along with their music, are organized as colored circles. The circle colors reflect the genre of the music and the bands with more fans and listeners are placed closer to the center of the screen. The user is able to select a circle and access the music in addition to biography, fans, and messages left for the band.

Sonarflow (2012) organizes a user's personal music collection into colored bubbles similarly to Hitlantis (2009). The artist bubbles can be pinch zoomed to view a deeper level including artist details and available songs. Double clicking a bubble will automatically generate a playlist and start playing instantly. Sonarflow also has a separate Spotify application that uses the same appearance to browse through online music. Similarly, Discovr (2011), Bloom.fm (2013), Musicbox (Lillie 2008), and Musicream (Goto & Goto 2005) map related artists into similar shapes and let users discover music accordingly. Discovr (2011) enables following of their favorite artists and the application will notify them when new content is available.

Liveplasma (2005) lets the user query for an artist, book, or movie and builds connections between related content as colorful bubbles and, in the case of music, starts playback instantly. The user interface also shows other albums from the artists in a sidebar and a music video from Youtube (2005) if available. The user has an option to solely play the searched artist or also include similar artists to the automatically generated playlist. The user is able to freely browse through the bubble space of artists relating to the initial search.

A cluster visualization experiment by Bossard et al. (2009) describes a visual music discovery interface using a lens and cake metaphor. The music is mapped to the user interface with small colored circles where colored sectors describe the musical style (i.e., cake metaphor). These circles are overlayed on top of large rings indicating the proximity of the circles to a center song (i.e., lens metaphor). The interface was evaluated with nine participants and compared to Sony's commercial SensMe (Sony SensMe 2014) interface. In the evaluation, the users were building new playlists, rating the songs selected for a playlist, and answering a questionnaire based on their experiences. The evaluated prototype outperformed state-of-the-art alternatives in terms of overall playlist quality, grouping of songs, and interest towards using the system again.

In Knees et al. (2006) the music collection is presented using a virtual landscape, a metaphor originally developed by Pampalk (2001). The landscape consists of islands, which are clusters of similar sounding music and the height corresponds to the density of the songs in the area. The user is able to listen to the sounds close to the listener's position in the virtual landscape. The landscape is further enriched with semantic information retrieved from the Internet. A similar type of mobile application is presented in Huber et al. (2012).

MusicGalaxy (Stober 2011) displays songs as stars (or alternatively album covers). The user can explore music by moving the focus of a primary fish-eye lens and zoom into different regions of interest. The neighbouring stars contain similar tracks to the focused area. Similar songs which are located further away due to projection errors (e.g., when using multiple attributes for similarity) are shown as distorted space, i.e., wormholes and are thus visible in the user interface. Musicgalaxy allows the advanced users to use different descriptors, such as rhythm, timbre and dynamics, to adjust the visualized similarity metrics and thus discover music more precicely. More interesting work regarding adaptive musical interfaces by Stober is found in (Stober 2011).

2.4.9 Music discovery using mood

Automatic mood recognition using different sensors is a difficult task. Different measurement options are discussed in Desmet (2003). In MIR systems, the user typically selects the mood with specific user interface elements. These elements include, e.g., buttons, emoticons, textual labels, or an interactive emotional space. Based on the user selection, the system commonly compares it with a database of mood metadata and returns a set of the best matching music.

Musicovery (2011) is a music discovery application roughly following Thayer's model (Kim, Schmidt, & Emelle 2008), where the user can select music according to mood, artists, tags, and genre. The mood can be selected by clicking on an x-y space, where y-axis has been mapped to energetic/calm music and the x-axis to dark/positive music (Figure 12).

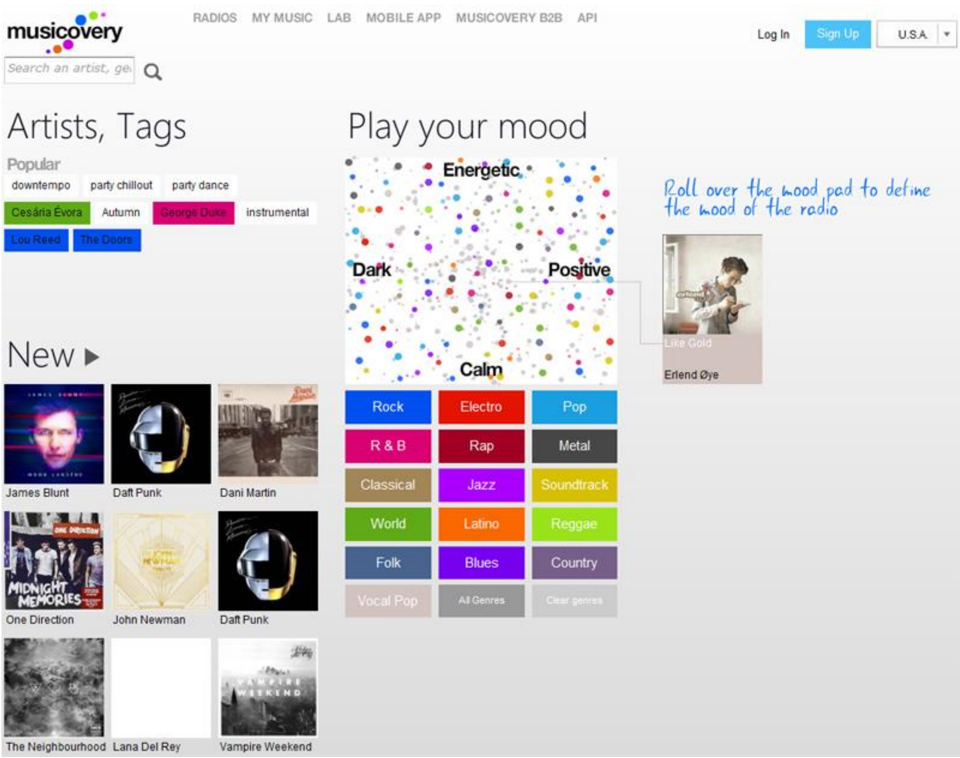

Figure 12 Musicovery (2011): Mood-based music discovery using x and y-planes to map the songs from dark to positive and calm to energetic.

The mapping is automatic, and does not require any effort from the user. When the user moves a cursor over the x-y space, preview clips of songs are played. When the user makes a selection, a playlist is automatically compiled and the user is able to change the playlist contents by changing decades and genres on a separate view. The changes are applied to the x-y space visualization.

In addition to 2D spaces, moods and emotions have been visualized in various ways including discs, color maps, icons, and vertical bars (Kim et al. 2009), and also emoticons, and tag clouds.

TheSixtyOne (2008) focuses on full screen album art and lets the user browse through popular and recent songs. The service also categorizes music based on moods (mellow, party, happy, trippy, crazy, smooth, sad, rocky, love, funny, remix, and covers). One special feature in the application is that while listening to a song and looking at the full screen album cover, artists' comments and small facts appear on the screen.

Songza (2007) combines mood and taste for music categorization and asks users questions while listening to music to learn more about when certain types of music is consumed. The service features user-curated playlists to match each mood and taste combination. Recently, Songza includes weather channel data as a new attribute for mood suggestions (Songza 2014).

With the Moodagent (2001) application, the user is able to adjust the balance between five mood bars: sensual, tender, joy, aggressive, and tempo. The resulting playlist is automatically generated based on the mood bar balance.

Stereomood (2012) represents mood tags with associated pictures with the user being able to make a selection for instant playback of music.

Approaches featuring colors representing moods include, e.g., the Color Player application (Voong 2007) with manual assignment of colored tracks based on their mood. Associating colors with musical genres has also been studied by, e.g., Holm et al. (2009) and Julià & Jorda (2009).

2.4.10 Discovering music using avatars

Castranova (2003, p. 4) defines an avatar as the "*representation of the self in a given physical environment*". Using avatars opens opportunities for reflecting the real or imagined identity of the user and can be used as an entertaining medium for interacting with the surrounding environment. Music can be used to relay messages regarding a person's identity (Rentflow & Gosling 2003). Avatar appearances can similarly reflect identities in virtual environments such as music discovery systems.

An interesting example application of mapping musical parameters to an avatar is found from Haro et al. (2010) and Bogdanov et al. (2013). The implemented prototype analyses music for over 60 low-level features from each song in a user's personal music collection. These features are used for building a set of semantic descriptors such as genre and mood. A subset of these features is used as attributes for the avatar character visualization in the user interface (Figure 13). The study had a strong algorithmic focus.

The prototype was evaluated with 11 participants using a questionnaire. The results conclude that avatars are able to provide a reliable but coarse visual representation of the user's musical preferences (Haro et al. 2010).

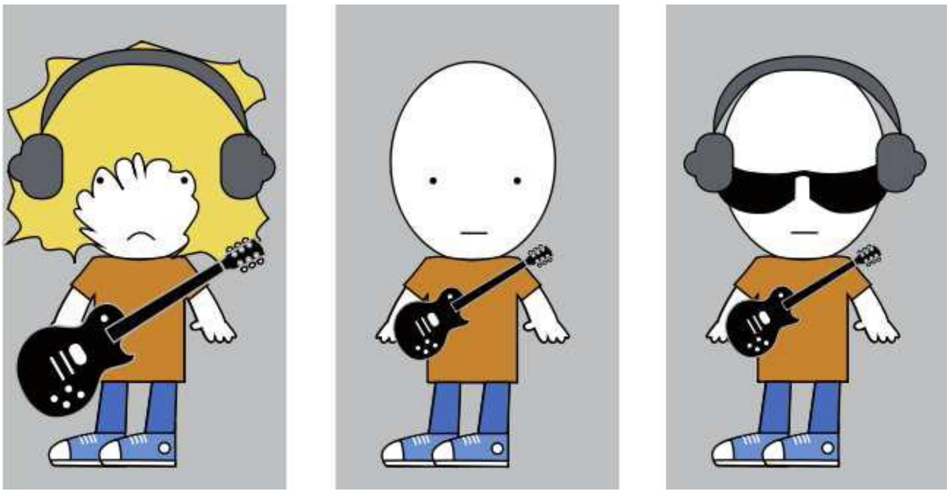

Figure 13 The Musical Avatar (Haro et al. 2010): Avatar examples for music discovery. (Used with the permission of the author)

Avatar design is addressed in more detail in Boberg et al. (2008) and other books devoted to the topic.

2.4.11 Audio- and gesture-based music discovery

Using spatial audio and gestures as an interface for controlling music is very common. The area of focus seems to be the use of such solutions in performances and musical compositions as novel interfaces. There are much fewer alternatives in the field of discovering existing music. Some examples are introduced in this section.

Keränen et al. (2009) describe the development and evaluation of a gravity sphere concept. Gravity sphere enables the user to explore a music collection by tilting the mobile device to different orientations. While tilting the device, preview clips of the songs are played back. The user is able to lock the orientation and a playlist featuring similar songs is automatically created. The evaluation results indicate that the interface is suitable for music exploration and the physical tilting of the device helps in memorizing the different directions associated with certain musical styles. The evaluation also indicates that cues are needed for the user to disregard previous assumptions regarding tilting gestures and their functionality towards music collections. Users additionally requested a feature for being able to manually organize music with the system and this is investigated using a bimanual interaction technique in Bergman et al. (2009). The bimanual interaction technique is a combination of touch input and physical tilting of the device for mapping music into different device orientations (Figure 14). The user evaluation shows that the proposed bimanual interaction is significantly faster to use than using the touch screen alone and was found pleasant to use to explore music mappings on a mobile device.

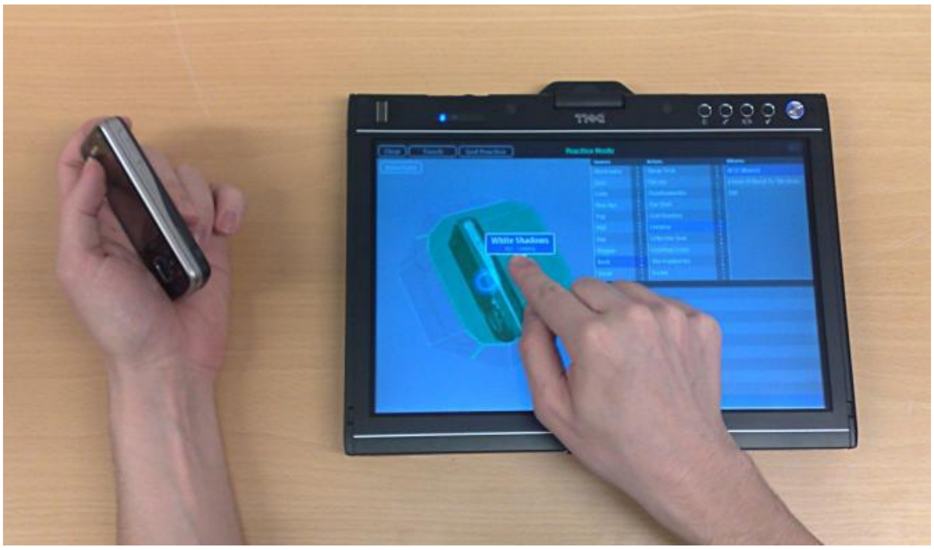

Figure 14 An example of audio and gesture based browsing and discovering songs bi-manually (Bergman et al. 2009, used with the permission of the author).

A study featured in Hiipakka & Lorho (2003) presents an interface where the user is able to browse a hierarchically arranged music collection, create playlists using a 5-way key, and receive feedback with audio only. The preliminary evaluation of the interface shows promising results in terms of most users being able to recognize the structure of the music collection after five minutes of use.

2.4.12 Summary of music discovery applications

This section has introduced a vast number of different applications and services for music discovery, categorizing them from the users' point of view. The introduced music discovery categories and application examples are summarized in Table 2.

The applications vary from performing textual searches to exploring artists using radial shapes. The recommendation types vary from handpicked music recommendations to fully automatic recommendations. There are many alternatives to choose from, matching the user's preferences for music discovery. Many of the services appear rather conservative and thus less playful or fun, and utilize tradional textual lists and basic visual elements as a basis of design. Most of the services do not fully support exploratory music discovery and expect a certain degree of musical knowledge from the user.

Music discovery category	Related application examples
Versatile services	Spotify, Deezer, Grooveshark
Radio-based	Pandora, Nokia MixRadio, 8tracks radio
Handpicked music	KCRW Music Mine, Stereogum, Beats music
Community-based	Soundcloud, The Hype Machine, Last.fm, Rdio, This is my jam, Shuffler.fm
Contextual	Soundwave, ContextPlayer, Foxtrot, Contextual recommendation prototype
Audio recognition-based	Shazam, SoundHound
Textual query-based	Youtube, Seevl, TasteKid
Using bubbles, radial and other geometrical shapes	Hitlantis, Sonarflow, Discovr, Bloom.fm, Musicbox, Musicream, Liveplasma, Cluster visualization experiment by Bossard et al., Islands of music, MusicGalaxy
Mood-based	Musicovery, TheSixtyOne, Songza, Moodagent, Stereomood, Color Player
Using avatars	The Musical Avatar
Audio- and gesture based	Gravity sphere, Bimanual technique for music mapping, Audio-only browsing of music collection

Table 2 Summary of introduced music discovery application categories and example applications.

2.5 User experience and playfulness in music discovery

This section first addresses the user experience and related research regarding music discovery. The role of interaction design is then briefly described. Lastly, playfulness and the use of stereotypes and visual metaphors are introduced and linked to music discovery.

2.5.1 User experience in music discovery

Hassenzahl et al. (2000) identify hedonic and ergonomic factors for determining a software's appeal. These factors include functionality, content, presentation, and interaction to create the overall user experience. Kuniavsky (2003) states that while the good user experience of a system can vary between individuals and tasks, a good approximation of the experience can be done by making the system functional, efficient, and desirable to its intended audience. Hassenzahl (2003) further discusses the relationship between the user and the product by building a model from both designer and user perspectives, where the designer is concentrating on the product features trying to relay intended product characteristics for the end user. The end user constructs an individual apparent product character, which leads to consequences – a judgment of the product's appeal. Product characteristics are divided into pragmatic *(*clear, supporting, useful, controllable) and hedonic (outstanding, impressive, exciting, interesting). The pragmatic and hedonic attributes together form a combination that formulates product characteristics.

User experience can change over time and, therefore, measuring long-term user experience is important. Karapanos et al. (2009) divide product adoption into three phases: orientation, incorporation, and identification. The orientation phase includes initial experiences which are judged based on excitement, frustration, and encountered learnability flaws. The incorporation phase emphasises the meaningfulness of a product for daily lives and incorporates long-term usability. The identification phase includes social and personal attributes including self-identify communication and differentiation. All three phases overlap in time but in the conducted study, experiences relating to the orientation phase drop sharpy after the first week of product usage. Attributes related to the incorporation and identification phases grow after the initial experiences (Karapanos et al. 2009). The long-term user experience is also summarized as follows: "*long-term UX is characterized as the overall summarized momentary experiences subjective evaluation of a product and its personal meaning after an extended period of usage*" (Kujala et al. 2013, p. 2). Longer-term usage and related user experience is understood in this thesis to include experiences beyond the first impressions and include a meaningful time of unsupervised use of the prototypes depending on the prototype complexity.

Wiklund-Engblom et al. (2009) study the relation of human needs and user experience. The six fundamental needs are connected to a variety of product attributes: autonomy (e.g., flexibility and expressing own opinions), relatedness (e.g., identifying and sharing), competence (e.g., usability and novelty), stimulation (e.g., emotional stimulation and creativity), influence (e.g., connectedness and communication), and security (e.g., reliability and familiarity). The design for different needs may include completely different requirements.

Another aspect of user experience, co-experience, has gained visibility recently. Battarbee (2004) proposes the use of real social and physical contexts of potential users in early phases of design. An understanding of the relevant user experiences can be obtained by observing how users boost experiences for each other, interpret, and respond to them.

This also allows the investigation of product interaction and new purposes for technology. Väänänen-Vainio-Mattila et al. (2010) expand the notion of social user experience by introducing drivers and hindrances based on a field trial from several social networking services. The most important positive drivers included self-expression, curiosity, reciprocity, and learning. The most common negative factors (i.e., hindrances) included suitability of the content and functionality, completeness of user networks, and trust & privacy.

In MIR, the key factor from the user's perspective is user happiness (Manning et al. 2008). It can be assumed that the delivery speed and relevance of the results is the key measure for user happiness. There have been alternatives to relevance including efficiency, satisfaction, and utility (Hu & Liu 2010). User happiness is often addressed by different means depending on the use case. Examples include completing a task, the amount of time saved, measuring return rate for users, speed of response, uncluttered user interface, and finding the wanted item (Manning et al. 2008).

Konstan et al. (2012) describe the transition of focus in recommender systems from algorithms to user experience. The user experience in such systems is described in terms of how the recommendations are delivered to the user and the interaction possibilities with the recommendations. The research is centered on recommender systems and their evaluations, whereas the holistic system and understanding different use cases for music discovery are not in the main focus.

As summarized in Holm (2012), relatively few empirical MIR user studies have been published that address new visual music discovery prototypes. Weigl & Guastavino (2011) summarize user studies in the music information retrieval literature. The summary states that a variety of algorithms have been developed to tackle various aspects of MIR but a full-featured system integrating these various capabilities is still missing. The evaluations in academic databases mainly involve evaluations of individual algorithmic approaches and their performance rather then overall user experiences, especially towards music discovery systems. So far, it is almost impossible to find any holistic user evaluations that would include longer-term user experience in novel music discovery systems. Additionally, the few existing evaluations do not regard the fun, entertainment, and playfulness sides of music discovery, which can be very important in explorative music discovery.

It is unlikely that there are no studies and experiments about the vast number of commercial systems for music discovery. The studies and results mainly remain company confidential for business reasons.

Some publications address understanding users in music discovery and categorization. Hu & Liu (2010) summarize the user studies in MIR. These user studies were described as focusing on identifying users' music information needs and the features used to describe them. In Schedl et al. (2013), many user evaluations in MIR are summarized and discussed. The approaches are stated to be still more system-based than user-centric. Three reasons are claimed to have prevented a major breakthrough in user-centric MIR: 1) "*general lack of research on user-centric systems*", 2) "*a lack of awareness of the limitations and usefulness of system-based evaluation*" , 3) "*the complexity and cost of evaluating user-centric systems*" (Schedl et al. 2013). In a study by Vignoli et al. (2004), genres are found as an effective tool for classifying unknown music on a coarse level but are not particularly good for finding known bands and popular songs. Users rated highly finding similar songs to a chosen song or artist in the study. An ethnographic study of searching and browsing techniques was performed by Cunningham et al. (2003). In the study, behaviors that are not well supported by current MIR systems were identified, e.g., the ability to sort search results by release date. A study of everyday encounters with novel songs (Cunningham et al. 2007) explores how individuals discover new music and proposes design implications for MIR systems. The key implications include providing related information to complement the music, laid-back searching (i.e., offline Internet queries are stored to be performed later when online) while capturing social interactions, and including richer metadata in audio fingerprinting. Users' needs and activities relating to music listening are studies by Arhippainen & Hickey (2011). The study introduces a wide range of listener and musician groups with clarifying use-case descriptions. Examples of identified listener groups include active listener, work listener, live listener, indie listener, and quality listener. The different groups can be used to help designing targeted music services.

2.5.2 Interaction design

Interaction design is important when designing services for music discovery. The user experience of a product can be improved using suitable interaction design methods. The difference between an engineering and interaction design approach has been described in Jones & Marsden (2006). A clearly defined problem is solved in engineering whereas in interaction design it is also important to form a view of the actual problem and then produce a solution.

According to Jones & Marsden (2006), interaction design is divided into three types of activities: understanding users, developing prototype designs, and user evaluations. Understanding users includes the biological features and limitations of a user. The evaluation phase introduces different options including testing with users (e.g., lab tests, field trials), testing without users (e.g., usefulness and analytical models for text entry using heuristic and expert based review). Very commonly, the development is iterative and new rounds and more detailed, focused evaluations are needed.

2.5.3 Playfulness

Korhonen et al. (2009) discuss understanding playful experiences through digital games. This related work states: "*that the behavior is pleasurable, rewarding or reinforcing and that the behavior is exaggerated, are important characteristics of playful experiences*." (Korhonen et al. 2009, p. 3) The paper defines an initial 20 categories for playful experiences and an extended list of 22 categories (Arrasvuori et al. 2011, p. 2) are listed in Table 3.

Experience	**Description**
Captivation	Forgetting one's surroundings
Challenge	Testing abilities in a demanding task
Competition	Contest with oneself or an opponent
Completion	Finishing a major task, closure
Control	Dominating, commanding, regulating
Cruelty	Causing mental or physical pain
Discovery	Finding something new or unknown
Eroticism	A sexually arousing experience
Exploration	Investigating an object or situation
Expression	Manifesting oneself creatively
Fantasy	An imagined experience
Fellowship	Friendship, communality or intimacy
Humor	Fun, joy, amusement, jokes, gags
Nurture	Taking care of oneself or others
Relaxation	Relief from bodily or mental work
Sensation	Excitement by stimulating senses
Simulation	An imitation of everyday life
Submission	Being part of a larger structure
Subversion	Breaking social rules and norms
Suffering	Experience of loss, frustration, anger
Sympathy	Sharing emotional feelings
Thrill	Excitement derived from risk, danger

Table 3 The PLEX Framework with 22 categories of playful experiences (Adapted from Arrasvuori et al. 2011).

In Arrasvuori et al. (2010), the playfulness in user experience is explored with personal mobile products. The most commonly appearing playful user experience categories in the trial were control, discovery, simulation, captivation, and humor. In this thesis, the playfulness in music discovery is defined using the following categories (but not limited to) from the PLEX-model (Table 3, highlighted items): captivation, discovery, exploration, and humor.These categories have worked as a baseline when designing different alternatives for playful music discovery.

2.5.4 Stereotypes and visual metaphors

Stereotypes in media have been visually characterized by Veletsianos (2010, p. 1): "*Humans draw on their stereotypic beliefs to make assumptions about others*" In the study, virtual characters in education were researched using pedagogical agents. The pedagogical agents created for nanotechnology and punk rock were stereotyped by the participants based on their non-verbal cues. The results show, e.g., that the punk rock agent was categorized to be more knowledgeable in the music domain compared to nanotechnology. In Veletsianos. (2010, p. 2) it is mentioned that "*First impressions frequently influence and guide subsequent interactions in real life*" (i.e., primacy effect). When interacting socially with unknown people, humans use visual cues to form their expectations guiding their following interaction (Brahnam 2001). The visual appearance of other people also has an effect on our attitudes and behavior (Haake et al. 2008). Visual stereotypes can be encountered, e.g, in photos, comics, and movies. Stylization enables the exaggeration and amplification of selected attributes commonly found in graphical media (Haake et al. 2008). The exaggeration and amplification of selected attributes can also be applied to music discovery concepts (e.g., **P3**, **P4**, **P7**, **P10**, **P11**). Further details regarding the psychology of stereotyping can be found in, e.g., Schneider (2005).

Visual metaphors are visual expressions of metaphorical thoughts or concepts (El Refaie 2003). El Refaie (2003, p. 78) states that, regarding metaphors; "*there is still little agreement among researchers even over basic terms and definitions*." The Oxford dictionary (Oxford 2014) define metaphor in an easily understandable form: "*A figure of speech in which a word or phrase is applied to an object or action to which it is not literally applicable*". Forceville (1994) has studied pictorical metaphors and defines a visual metaphor as being the replacement of an expected visual element by an unexpected one.

Averbukh et al (2007) propose a set of criteria for metaphor generation including:

- object property similarity (source and target domains)
- ability for a visual presentation of an object (source domain)
- object recognition (source domain)
- number of interrelations between objects (source domain)

Using metaphors and supporting existing user models of a phenomenon and processes (i.e., representative cognitive structures) are the building blocks of successful HCI and visualization (Averbukh et al. 2007). The use of stereotypes and visual metaphors offer interesting opportunities in the field of music discovery for visually guiding user interactions and offering fun and entertaining experiences.

2.6 Summary and research gaps

This chapter addressed the related work for music discovery applications. The chapter began with the introduction of the analysis of musical files to retrieve metadata to be used in music discovery applications. Common low-level and high-level content-based methods were described with practical examples to understand the basis of automatic metadata generation. This was followed by other relevant sources about musical metadata. The next section addressed the most commonly used musical metadata in music discovery applications. Music recommendation systems are very important in music discovery. Different types of music recommendation methods were classified and summarized with their benefits and drawbacks. The chapter then proceeded to music discovery by introducing a state-of-the-art review of current music discovery applications and services. Finally, the user experience in music discovery was discussed.

Most music discovery services are commercial and, therefore, there are not many academically published user evaluation results available to help in the design or understanding of their overall user experience. There are a small number of evaluation results available for academic music visualization prototypes but the prototypes are often neither fun, playful, nor entertaining. The focus of these prototypes is mainly in music recommendation algorithm performance or visualization in terms of encoding the data. Music discovery and the user experience of such systems are not their focus. Additionally, there were no studies found including longer-term user experience for playful music discovery prototypes.

The thesis contributes to the fields of HCI and MIR with ten different concepts and their prototypes for music discovery and the evaluations of each of them with users. Several prototype evaluations include longer-term usage data beyond the first impressions. The evaluation results describe the user experience of such systems in detail and lead to the formulation of design implications to help build better music discovery focused services.

3. Research approach and methods

This chapter describes the reseach approach and methods used to study the novel music discovery prototypes. The work conducted in Nokia Research Center (NRC) during 2006-2014 consisted of separate parallel streams involving audio analysis algorithms and design for music discovery. The original goal of the music discovery stream was to design novel concepts and prototypes and learn from users how to design systems for easy, fun, and entertaining music discovery towards the "ultimate" music experience.

This chapter first describes the research approach and summarizes the research methods used. The three individual research phases are then described in detail with short prototype descriptions, research methods, and trial practicalities.

3.1 Research approach

The research of music discovery applications was conducted using a human-centered and constructive design-science approach (Mao et al. 2005).

Abras et al. (2004, p. 1) describe user-centered design as: "*User-centered design (UCD) is a broad term to describe design processes in which end-users influence how a design takes shape. It is both a broad philosophy and variety of methods. There is a spectrum of ways in which users are involved in UCD but the important concept is that users are involved one way or another. For example, some types of UCD consult users about their needs and involve them at specific times during the design process; typically during requirements gathering and usability testing. At the opposite end of the spectrum there are UCD methods in which users have a deep impact on the design by being involved as partners with designers throughout the design process*". The design principles in Jokela et al. (2003) further characterize user-centered design based on (currently deprecated) ISO 13407 including active user involvement, appropriate allocation of functions between users and technology, design solution iteration, and multi-disciplinary design.

A newer defition replaces user-centered design with human-centered design. The definition of human-centered design is found from ISO 9241-210 (2010): "*approach to systems design and development that aims to make interactive systems more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques*". Human-centered design emphasizes a wider spectrum of stakeholders, beyond users, compared to user-centered design.

Hevner et al. (2004) describe the design-science research to include knowledge and understanding of a problem domain. Knowledge is achieved through building and evaluating innovative and purposeful artifacts. Artifacts refer, in this case, to music discovery prototypes. The gathered knowledge can be used in the design, realization (construction), and improvement of existing artifacts (Van Aken 2004).

Human-centered design supports a design-science approach well in understanding the user needs in a specific domain in order to aid the artifact (i.e., concept and prototype) design and evaluation. In the music discovery domain, it is important to understand different music consumption habits and to gain user feedback on the designed discovery methods.

The research goal was to design fun and entertaining solutions for music discovery, evaluate the solutions with users, and derive design implications for such applications. All of the different aspects of interaction design were used in research: understanding the users, prototype design, and evaluation (Jones & Marsden 2006). The focus in this thesis was that of prototype design and evaluation. The prototype evaluations gave insight of user needs for music discovery.

3.2 Summary of research methods

The research of music discovery prototypes is divided into three phases: initial concept design phase, playful concept exploration phase, and iterative concept design phase. Each prototype was evaluated with Finnish active music listeners gathering qualitative and quantitative data.

Active music listeners are potential users of music discovery systems and were, therefore, selected as the main target group. There was no clear age limitation to take part in the trials but the research practicalities (i.e., gathering volunteers for trials) attracted mainly 20-40 year old participants. The participants were recruited in one or more of the following ways depending on the trial: mailing lists, advertisement boards, personal recruitment and social media. Participants were rewarded for their participation with a small gift, most commonly movie tickets. Altogether, 149 people participated in the studies. 64% of the participants were male, and 36% female.

A short summary of the performed user studies and their linkage to the research questions is shown in Table 4.

Publication	Time	Location	Users (N)	Data gathering methods	RQ
P1	2H/2007	Tampere and Helsinki, Finland	42	Semi-structured interview, questionnaire	RQ1, RQ3
P2-P7	1H/2010	Tampere, Finland	40	Observation, semi-structured interview, questionnaire	RQ1, RQ3
P8	1H/2010	Tampere, Finland	40	Observation, semi-structured interview, questionnaire	RQ1, RQ2, RQ3
P9	2H/2012	Tampere, Finland	15	Observation, semi-structured interview, questionnaire	RQ1, RQ2, RQ3
P10	1H/2012	Tampere, Finland	30	Observation, semi-structured interview, questionnaire	RQ1, RQ2, RQ3
P11	2H/2013	Tampere, Finland	22	Observation, semi-structured interview, questionnaire	RQ1, RQ2, RQ3

Table 4 Performed user trials and the linkage to research questions.

The qualitative data was gathered to find out in-depth information on ‘why’ and ‘how’ questions, which are not easy to address solely using questionnaires or observation. Rich qualitative data aids the formulation of future directions and design implications.

The qualitative data was analysed using a qualitative content analysis method. In many of the evaluations, the qualitative content analysis method followed the grounded theory methodology (Strauss & Corbin 1994). The grounded theory methodology is especially helpful when the current theories regarding the phenomenon are either inadequate, nonexistent, or the study investigates a certain process (Creswell 2008).

The qualitative data collection included recorded interviews and note taking. The recorded interviews were transcribed and the notes were used to help the coding and categorization phase. The notes also included observations from the interview session not accessible from the audio recordings. Biography data from the users were also used in coding the data. From the data collected, the key points are marked with series of codes, which were extracted from the transcribed text. The codes were grouped and further iterated towards larger categories, which are the basis for the creation of a theory.

Open coding and constant comparative data analysis were mixed and used simultaneously. The data was divided into segments and then reviewed for common categories or themes. Additionally, after each interview the data was coded based on the previous interviews in mind for coding improvement and theory generation (Glaser 1992).

The categories were further examined for properties that characterize each category. In general, open coding is a process of reducing the data to a small set of themes that appear to describe the phenomenon under investigation (Glaser 1992). Tables or concept maps were not used to represent emerging categories, thus constant comparative data analysis helped in finding suitable coding and categories.

This was followed by analyzing and reporting the results (theory). Theory can be seen as the results reported in the individual publications, including the conclusion and design implications.

Quantitative data was gathered to perform statistical analysis supporting the validity of the results, summaries of prototype performance in various aspects, and giving perspective to the interview comments in terms of numerical valuation.

The quantitative data for the prototypes in playful concept exploration and iterative concept design phases was gathered using a seven-point Likert scale and the results were visualized using boxplots (**P2-P11**). The seven-point Likert scale was chosen over five-point to give more granularity to the answers regarding experiences with the visual music discovery prototypes. Statistical analysis was performed in order to find out whether the answers depend of, e.g., gender or age. Chi-Square was initially used to investigate the problem. If the conditions for using the Chi-Square test were not met, Spearman's correlation and Wilcoxon's signed rank tests were calculated. The requirements for using Chi-Square include that no more than 20% of the counts (i.e., response options) have less than five answers and all counts have at least one. This condition was not commonly met with the answers to our UX questionnaires. Instead, Spearman's correlation was calculated to address the issue. Wilcoxon's signed-rand test was used to investigate the differences between the answers initially and after longer-term usage.

Each research phase is next described with further details on the used methods and practicalities.

3.3 Initial concept design phase

The research of music discovery systems was inspired by the development of an in-house content-based music recommendation algorithm. Lehtiniemi & Seppänen (2007) presents the evaluation results of the first algorithm version. The positive experiences of the algorithm performance led to the implementation of a mobile end-to-end system for on-demand music streaming and music recommendations. The first music discovery prototype was called SuperMusic (**P1**). At the time of development, Spotify or other streaming mobile music services did not yet exist and thus SuperMusic was considered a novel music discovery prototype (Figure 15).

Figure 15 SuperMusic example screen featuring from left to right: starting screen, options menu for main features, recommendation playlist, recommendation inbox, and a voting dialog for the audible song.

The main features of SuperMusic consisted of streaming music to a Symbian mobile device and both active and passive music discovery. The user was able to search for music from an online database of over 200,000 songs, receive contextual recommendations, and receive recommendations from friends using the system. User feedback was collected to further enhance the contextual recommendations.

SuperMusic was evaluated with 42 Finnish users (26 male, 16 female) during November-December 2007 to find out how users perceive the overall concept and the implemented prototype. The users were active music listeners with an average age of 30. Music was listened to every day by 88% of the participants and a mobile device was used for music listening, on average, in 32% of the cases. Almost half the participants played an instrument as a hobby and had experience with other music recommendation systems. All the participants were Nokia employees and technically oriented.

The evaluation consisted of a five-week trial period using a Nokia N95 device. During the trial period, weekly top listener charts were published to motivate the users to be active. The trial period was followed by a 1-2 hour semi-structured interview session including a questionnaire to collect quantitative data. The users were rewarded with movie tickets and the top three most active listeners received an additional reward. A qualitative content analysis method was used to process the interview results (Zhang & Wildemuth 2009, Lacey & Luff 2001). In Patton (2002, p. 453) qualitative content analysis is defined as: "*any qualitative data reduction and sense-making effort that takes a volume of qualitative material and attempts to identify core consistencies and meanings*". The SuperMusic interview data was first transcribed into textual format. The data was then browsed through, categorized, and interpreted. Finally, the data was summarized and reported. Additionally, lightweight statistical analysis was performed to the questionnaire data.

The results firstly drove the research into developing a second version of the prototype called SuperMusic v2. The new version included a number of improvements based on the findings in **P1** including, e.g., one-click play, dynamic top-lists and charts for easy music discovery, improved search functionality, and genre icons described in (Holm 2012). The recommendation algorithm was further improved and textual genre metadata was included to filter the recommendations. The new version was evaluated with the development team and selected others within the company but the results have not been published. The experiences with text-based versions of SuperMusic steered the research towards finding fun, entertaining, and visual ways for exploring large music collections.

3.4 Playful concept exploration phase

The playful concept exploration phase started with sketching a number of ideas that could make the music discovery fun, playful and thus an entertaining experience. Examples of initial sketches for visual music discovery by the author are shown in Figure 16.

Each sketch illustrated (in Figure 16) consisted of several screens to show the basic operational flow of the music discovery approaches. The initial ideas were analysed with a small group of company-internal specialists and external programmers and user interface designers.

Figure 16 Examples of initial concept sketches for visual music discovery

The ideas were further iterated in addition to generating new ideas. After two full-day iteration sessions a final set of six ideas were selected. The main selection criteria were: playfulness, novelty, ease of implementation, and how well the concept supported the existing recommendation back-end in order to avoid excessive algorithm development. The six selected music discovery concepts were then fine-tuned to be implemented as functional prototypes (Figure 17).

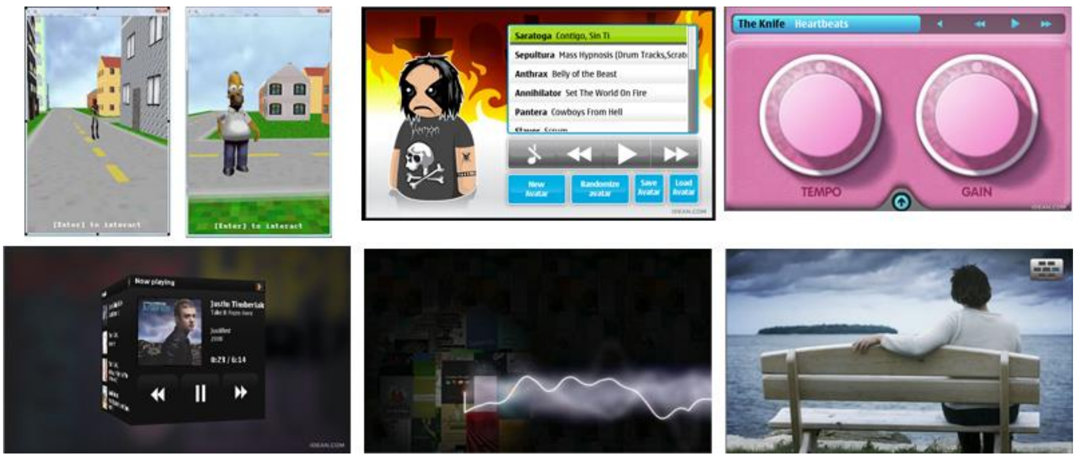

Figure 17 Concepts for playful exploration with 'working names' from top left to bottom right: Virtual world (P2), Avatar (P3), Potentiometers (P4), Cube (P5), Cover art space (P6), and Mood pictures (P7).

The implemented prototypes included (from top left to bottom right in Figure 17): Virtual world (**P2**), Avatar (**P3**), Potentiometers (**P4**), Cube (**P5**), Cover art space (**P6**), and Mood pictures (**P7**).

A virtual world was utilized in (**P2**) where users were able to move freely and interact with objects and characters. Different sections of the virtual word were visualized to match a certain genre and characters recommending music from those genres were found there. The characters were visualized to present the music that they are recommending (e.g., cartoon music in Figure 17). Recommended music could be added to a playlist and listened to while exploring the world.

Avatar characters were used to discover music in **P3**. The user was able to change the avatar's appearance by swiping the head, torso and background image to match their musical preferences, e.g., heavy metal (Figure 17). A playlist was then automatically built based on the selections. Each avatar part is associated with a set of typical songs for that genre and these songs are used as seeds by the music recommendation algorithm for generating the playlist. The avatar prototype enables the user to build playlists matching the visual appearance of the avatar and enables a fast access for adjusting genre balances in the playlist by mixing the avatar parts. The prototype implementation included five major genres to select from and combination favourites could be saved for quick access later.

Potentiometers were used for music discovery in **P4**. Firstly, a default visualization was shown to the user with similar tempo and gain potentiometers (as in Figure 17). The user was able to bring up a genre selection menu by clicking the blue arrow in the lower part of the screen. The prototype implementation included five major genres to choose from and the whole prototype appearance changed to match the genre selection, e.g., pop music in the example of Figure 17. Within a selected genre the user was able to adjust the tempo and gain potentiometers and hear a representative preview song with the corresponding settings. When the user was satisfied with the musical style, a playlist could be generated seeding the recommendation algorithm with the preview song. In the playlist view the user was able to further expand the listening experience by using any of the playlist songs as a basis for a new playlist.

A cube metaphor was used for music discovery in **P5**. When starting the prototype, a random cube was displayed (Figure 17). The music player controls, a recommended playlist, artist albums, and simulated artist information, biography and friends were mapped onto the faces of the cube. The user was able to rotate the cube and access the different features. By touching outside the cube (i.e., the blurred background) the cube would minimize to a "cube world" grid of cubes. The small cubes formed genre names with different colors and touching a small cube from those areas would maximize that cube and start playing music from that genre using a seed song associated to the selected cube.

Discovering music by moving a finger in the album cover space and listening to preview clips was introduced in **P6**. The music in the album cover space was divided into six sections (i.e., genres) and each section included a set of typical songs from that genre as preview clips. While moving a finger around the screen the preview songs change accordingly. For each section, a preview song was randomized when the user moved his finger to that area (Figure 17). When a suitable type of music was found and the finger was released, a playlist was automatically generated seeding the recommendation algorithm with the preview song and the application moved to a playlist view with music player controls. The user was able to listen to the playlist and additionally use any of the songs as a seed for a new playlist. The user was able to go back to the album cover space any time to browse for new music.

Selecting music using mood pictures was used in **P7**. The prototype starts with a randomized preset picture representing a mood or an emotional state and associated music startes to play (Figure 17). The user was able to browse through the seven built-in mood pictures and when changing the picture by swiping the display to different directions the music was also changed. The mood pictures were enhanced with small animated regions, e.g., overlayed raindrops in Figure 17. The button in the top-right corner enables the user to access a quick selection menu of different mood pictures and additionally shows the user their position in the current menu hierarchy. By tapping the mood picture, the user was able to access the currently audible playlist and music player controls. In addition to the generated playlist, similar songs visualized with album covers next to the playlist were proposed. The user was able to place a finger on top of an album cover to preview the recommendations and resume playback of the current song by releasing their finger. Preferred items could be dragged to a preferred position within the playlist. Clicking outside the player and playlist controls would fade away the menus and show solely the mood picture.

After finalizing the prototypes (**P2-P7**) a user evaluation was arranged with 40 Finnish participants (31 male, 9 female) in 2010. The participants were of varying age groups from 12 to 55 years old with the biggest concentration in 30-40 year-olds and the average age of the participants being 34.6. Most of the participants were active music listeners. 63% of the participants were university graduates and the rest had mixed educational backgrounds. Half the participants stated music listening as being their hobby and, similarly, half the participants played an instrument. Average music listening time for each day was estimated to be slightly over 2 hours for each participant. Music was mostly listened to at home and in a car (55% and 45% respectively) and the most popular listening mediums were radio, CD and MP3 players including mobile phones. 43% of the participants stated that they actively search for new music and that their typical sources for music discovery included the Internet (especially Youtube), friends, magazines, and Spotify.

The initial research questions covered general prototype usage experience and further development, including entertainment value and user experience after a longer-term usage period. The suitability of replacing or complementing current music applications was also investigated. The trial started with filling a background questionnaire and proceeded to testing the prototypes, complemented with a semi-structured interview session and questionnaires. All participants were observed while testing the prototypes in a randomized order. The testing on-site was done using a HP Pavilion tx2500 touch-screen laptop. After testing each of the prototypes for 5-15 minutes the users were interviewed regarding their experiences followed by a short questionnaire covering the quality aspects of each prototype. No instructions were given regarding the prototype usage and the users could explore them freely. The interview sessions included both hedonic and pragmatic aspects of the prototype usage and the quality aspects were measured using a 7-point Likert scale.

After the prototype testing and an interview session, the users were encouraged to use the online version of the prototypes (**P3-P7**) at home for three weeks. After the three-week usage period, the users filled in another questionnaire regarding the quality aspects of the prototypes. The longer-term usage period did not include prototype **P2** due to its local implementation.

A qualitative content analysis method was used to process the interview results (Zhang & Wildemuth 2009, Lacey & Luff 2001) and statistical analysis was performed on the questionnaire data. The questionnaire data was used to better understand the interview comments with the help of the numerical assessment of similar topics.

3.5 Iterative concept design phase

The results from **P2-P8** encouraged further development of the ideas. A different view on music-specific social interaction was experimented on in **P9** (MyTerritory). The earlier results showed a big conceptual potential in **P3** (Avatars) and **P7** (Mood pictures) with a fundamental re-design. The results were used to enable new experiences using the original concepts of using mood pictures and avatars for music discovery (**P10** and **P11**).

3.5.1 MyTerritory – A world of music

To incorporate outdoor environment and collaborative game-like experiences to music discovery, a prototype called “MyTerritory – A world of music” was created (**P9**). The prototype enabled the users to conquer physical areas by circulating them. Once an area was conquered, the user was able to associate a song with the area and then the area was visualized accordingly (Figure 18, middle).

Figure 18 Example screens from the MyTerritory prototype: conquered territories (middle) and music selection based on routing (right).

The area could later be captured by circulating the area more times or being more active while circulating. An in-house activity-detection algorithm was experimented with in the application for measuring the user's effort while conquering an area. Each conqueror was able to associate a song with the newly conquered area and eventually the area featureed a playlist of the associated songs. In addition to conquering areas the users were able to discover music in three different ways: playing a territory with associated tracks, selecting an area from the map view to play all songs within that area, and creating a route and playing all songs along the route (Figure 18, right). The application included a music player with playlist management options.

The prototype was evaluated with 15 Finnish participants (8 male, 7 female) during 2012. All the participants were active music listeners and the average age was 27. The users mostly included graduate students (66%) and almost half the participants worked in the ICT field. All users had experience in more than one online music service. The evaluation consisted of a field trial, a semi-structured interview, and a questionnaire. In the field trial Nokia E7 smartphones were used. Before the actual trial session, a brief introduction was given to the overall concept. This was followed by performing guided tasks using the prototype. After the field trial, the users were interviewed regarding their gaming experiences, and the overall concept and its further development possibilities. Additionally, observations from the field trial were clarified in the interview sessions. The evaluation results were processed using qualitative data analysis using a grounded theory methodology (Strauss & Corbin 1994). Questionnaire data was statistically analyzed. Additionally, the questionnaire data was used to better understand the interview comments with the help of the numerical assessment of similar topics.

3.5.2 Collaborative mood pictures (MoodPic)

The feedback from the first mood picture prototype (**P7**) encouraged refining the concept and re-building the prototype. Firstly, the preset mood pictures were re-designed to involve people (Figure 19, left). Secondly, the whole concept was changed to feature collaborative playlist generation and the ability to use users' own pictures as a basis for mood playlists (Figure 19, right). Each user was able to associate songs from the internal music catalog to any picture in the system and thus collaboratively create a playlist matching the mood picture. Every song addition to the playlists was complemented with the name of the user who contributed it. Each preset mood picture had five songs already selected for quick access to music but they could also be extended with users' additions. When listening to a playlist, the mood picture was displayed as a full screen background image to enhance the experience (Figure 19, middle). Thidly, the users were able to customize their starting screen with their favorite mood picture playlists (Figure 19, left).

Figure 19 Example screens from the Collaborative mood pictures (MoodPic) prototype (P10) including starting screen with quick shortcuts (left), listening a mood playlist (middle), and mood playlist gallery (right).

The prototype was implemented using HTML5 and separate desktop and mobile layouts were designed. The prototype was evaluated by 30 Finnish participants (16 male, 14 female) during 2012. All the participants were active music listeners. The participants' ages varied from 19 to 56 with an average of 29. The initial research questions included themes regarding collaborative playlist creation for music discovery, overall usage experience and further development, and design implications for music service developers.

The users were able to try the prototype freely for 2-5 days before the interview session. No instructions for the use were given to the users. Users were assigned simple tasks via email to make sure that they had tried all features of the prototype, e.g., creating a mood picture playlist and recommending songs to others' playlists. These tasks were discussed in the semi-structured interview sessions arranged after the trial period. Each two-hour interview session was divided into three parts: 1) walkthough of the system with user's own account and discussion about different features and the created content, 2) general semi-structured interview regarding the concept and the user experience of the system, 3) filling in a questionnaire regarding the user experience of the system. A qualitative content analysis method was used to process the interview results using a grounded theory methodology (Strauss & Corbin 1994). The questionnaire data was used to better understand the interview comments with the help of the numerical assessment of similar topics.

3.5.3 Adaptive Avatar music experience

The results from the first avatar prototype (**P3**) showed potential in the use of avatars in music recommendation. The prototype concept was re-used and further developed into a completely new prototype (**P11**).

In the new prototype, the avatar represents the user's recent music listening history and constantly evolves according to the user's music consumption (Figure 20).

Figure 20 An example starting screen from the Adaptive Avatar prototype (left) and the collection of built-in avatars (right).

Users will first encounter a naked avatar. After each listened song from the internal music database, the avatar is populated with an element (e.g., clothing or accessories) representing the genre and the era of the listened song. The listened song is associated with the added element. Depending on the music consumed, the avatar will have a different appearance – either consistent to a certain style or fully mixed. When the avatar is fully populated (i.e. dressed and equipped), the new songs will either change the avatar's appearance or populate playlists associated with the individual elements. The user is able to listen to the created avatar with one click or listen and explore the individual avatar's parts. Users are able to see and listen to others' avatars in the system. The prototype enables users to send avatar parts with the associated songs as visual recommendations and store their favorite parts to a wardrobe for later access. In Figure 20, right, there is the complete avatar selection visualizing 28 different avatar characters (consisting of appromately 10 elements each) representing different genres and eras. The user's avatar can be either one of these, if the music is listened consistently, or a complete mixture of the avatar parts having an almost endless amount of available combinations.

The prototype was evaluated by 22 Finnish participants (15 male, 7 female) during 2013. All the participants were active music listeners and most of the participants were actively listening to music with mobile devices. The average age of the participants was 30.4. There was a big variation in mobile music listening habits. Mobile music was consumed mostly while working, exercising, and travelling.

The user trial started with an introductory session where basic concept information was given, an avatar building task was performed, and background information from the users was collected. All of the preset avatars in the system were printed out and cut into three pieces. The users performed an avatar building task in the introductory session stating their musical taste, combining printed avatar parts to form a complete character.

The users were then able to freely use the prorotype for 1-3 weeks and the system automatically gathered a listening log to monitor the user activity in the system. During the trial period, 1849 songs were listened in total, which divide into 84 songs, on average, for a user. After the trial pediod, a semi-structured interview session was arranged to find answers to initial research questions. The research question themes included the overall user experience of the system and support for music discovery, visualization of recent music listening history, avatar designs for visualizing genres and eras, and how well avatars work as social enablers in a music discovery system. Quantitative data was collected using a questionnaire and the system usage log. A qualitative content analysis method was used to process the interview results using a grounded theory methodology (Strauss & Corbin 1994). The questionnaire data was used to better understand the interview comments with the help of the numerical assessment of similar topics.

4. Results

This chapter summarizes the key results regarding the research questions listed in Section 1.3. The main aim of this research was to understand the user experience of novel music discovery concepts through prototype implementation and user evaluations, according to the design-science research approach (see Section 3.1).

4.1 What is the user experience of novel music discovery prototypes?

The following sub-sections describe the positive and negative user experience highlights of each music discovery system.

4.1.1 Streaming mobile music client with recommendations for music discovery

SuperMusic (**P1**), the streaming mobile music service was seen as a great overall concept with 97% user satisfaction and 95% of the participants would have liked to continue using the system after the 5-week trial period. Altough the service suffered from technical streaming related issues and user interface related inconsistencies, the ability to stream music to a mobile device with music recommendations in 2007 resulted in an active participation for the trial. During the trial period, users tried over 25 000 songs from the service, which is, on average, 17 songs each day per trial user. The amount of songs listened to shows that the novelty value surpassed the usability and technical problems of the prototype system. Despite the problems, 73% of the users considered the prototype easy to use. The users voted 63% of the similar songs recommendations to be good and 75% of the users considered the similar sounding music recommendations to be a very important feature of the prototype system. Also, 63% of the users considered displaying the recommendation criteria important. SuperMusic enabled sending music recommendations (i.e., sending a link to a song from the music catalog) to other users in the system. This type of social interaction was appreciated and 90% considered social interaction in such music service beneficial.

The user experience of SuperMusic suffered from user interface inconsistencies and unclear terminology such as “situations”. Some features, e.g., recommendations from friends were hard to find if the user missed the automatic notification dialog. Users were not able to fully utilize the large online music catalog due to the limited ways of accessing the music. The contextual recommendations were not working well, therefore, the practical alternative was to first know a seed song to start with and then retrieve similar sounding playlists from the service. Top-lists sent via email during the field trial inspired the users to try the most popular songs and made music discovery easier. Users could vote for the recommendations using a simple Good/Bad dialog. A user's effort in voting did not result in any change in the current playlist, which resulted in a bad user experience and the feeling of not being in control over the recommendations. Song recommendations from other users were sometimes confusing due to the lack of ability to send links to music with any description, e.g., not knowing why they were sent.

4.1.2 A virtual world for music discovery

The virtual world prototype for music discovery (**P2**) was seen as a promising concept but the implementation failed to support the user needs and preferences for music discovery. Altough the concept was seen as novel and easy to use by most of the users, the relationship between the implemented concept and music discovery was unclear. The prototype was even initially confused with a first person shooter game. Almost half of the participants saw the prototype being more fun than traditional music players but the music discovery was way too time consuming and cumbersome for the users having to run around in the virtual world and encounter mostly poorly visualized characters to get music recommendations. The visual appearance of the prototype was seen old-fashioned or simply bad, which did not boost the overall experience. The prototype was designed within a student project and the outcome, especially regarding the character design, was not according to the specifications. The prototype was said to have too few features and not enough things to do in the virtual environment. Having music recommendations as the sole activity in a virtual world was not the right approach. Most of the participants were not interested in using the prototype in the future, even to complement other music players. 88% of the users saw the application not suitable to their current music consumption habits and several participants commented on not being interested in the application even if the implementation were improved.

4.1.3 Avatars for music discovery

The first prototype that used avatars for music discovery by changing the avatar appearance to create playlists (**P3**) received a relatively positive first impression with 65% of the participants stating the concept to be easy, novel and fun. 18% of the participants were generally positive but the use of cartoon-like avatar characters was seen to target younger audiences by several participants. For 13% of the participants the first impression was confusing in terms of how the application works and it resulted in a bad initial user experience. However, the participants were able to use the system after a few minutes of exploration. The automatic playlist generation based on the avatar appearance was said to work well for people that are open to many types of music but not well suited for listening to a specific type of music. The strong use of stereotypes in the avatar design was said to increase the fun and entertainment value of the concept. While most users were able to recognize the musical styles that the avatars presented, some participants were not able to recognize all the avatar characters correctly and the resulting music decreased their overall user experience. The prototype had very limited user control in playlist generation and thus the comment is relevant and justified but does not directly link to the overall use of avatars for music discovery. Using avatars for music discovery was generally seen as inspiring, innovative and surprising. Receiving music recommendations based on the customizable avatar appearance was seen as a good idea. The concept was said to inspire users to explore music and provided quick access to different types of music and cross-genre playlists in comparison with traditional music player applications. In longer-term use, the prototype was seen limited in terms of different features and the overall grade for the system was lower compared to the initial impressions. The prototype was targeted to music discovery and lacked basic music searching features, which was a major shortcoming for many of the participants especially in longer-term use. Only 10% of the participants would use the prototype as such as their only music application but 65% would use it to complement other music applications.

The second avatar prototype used avatars for music discovery in a different way. In (**P11**), the avatar combinations were not used for creating playlists as such but the user's recent music consumption changed the avatar appearance accordingly and associated the avatar parts with listened music. The new approach was very well perceived. The service was said to be fun, interesting, novel, offering a new music listening experience, and providing a novel way to share music. The users were highly interested in their constantly evolving avatar character and different, partly unexpected, avatar creation strategies were discovered that changed the way users interact with a music service. The evolving avatars were seen to be even more interesting than the music offered by the service for some of the users. The majority of the users considered the approach of visualizing recent music listening history instead of the whole history to be a good idea. Most of the users who did not agree with the recent music listening visualizations wanted either an additional avatar for the whole history or the ability to hide certain songs from the avatar. Most users were interested in seeing and listening to other users' avatars in the system. Others' avatars were said to give a good overall picture of a person's musical taste and work as a tool for music discovery and self-expression. Most users wanted to express their musical taste and recent music listening using their avatar. Users were not able to see who viewed or listened to their avatar in the system, which was a shortcoming for some of the users. The use of adaptive avatars encouraged over half of the users to discover music more broadly compared to their own initial evaluation regarding their musical taste. The results partially reflect the longer-term use of the prototype since the system was used for 1-3 weeks before the interview session.

4.1.4 Potentiometers for music discovery

Selecting a genre and fine-tuning the tempo and gain –potentiometers to retrieve a playlist was the core of **P4**. The use of the prototype resulted in a very positive first experience for 60% of the users. The use of potentiometers for building a playlist was generally seen as easy, fun, exciting, different, and novel. Most participants liked the idea of genre selection affecting the whole application appearance. The potentiometers offered the ability to fine-tune the music recommendations within a genre and that was an especially appreciated feature. The tempo filter for playlist generation was seen useful and relevant but the gain (i.e., intensity) potentiometer was unclear for several participants. Most users understood the gain potentiometer function quickly while exploring with the settings but the term was not previously familiar for some. The prototype had some interaction design related issues that affected the user experience for some participants. Generally, the prototype was found easy to use and the interaction to be good for 90% of the users. In longer-term use, the prototype was found slightly too simple and the lack of a text-based search was a major shortcoming for longer-term active use. As a concept, the potentiometers work well for filtering playlists and 25% of the users suggested adding more (or customizable) potentiometers to adjust the playlist contents further.

4.1.5 Cubes for music discovery

Mapping musical information to the faces of a cube was used for music discovery in **P5**. The common first impression in using a cube for music discovery and mapping its faces with different musical features was that of fun but at the same time slightly confusing. The cube as an interface for music discovery was seen to be interesting and playful with a lot of potential. Mapping musical information to the different faces of a cube was seen to be good, especially when used in the mobile domain for saving valuable screen estate. Users highlighted the ability to view two faces simultaneously in the cube metaphor (Figure 17, lower left). The cube analogy was said to be easy to understand. The current implementation caused problems for many users in terms of interaction with the cube and thus offered a bad user experience. The interaction between the maximized cube and the selection of the cube (cube world) was not clear for many and made understanding the overall concept difficult. The feedback relating to the concept included some variation between the users and the overall grade was not admirable. In longer-term use the cube prototype was seen slightly too simple and the lack of a text-based search was seen as a major shortcoming.

4.1.6 Album cover space for music discovery

The album cover space -prototype enables the user to move a finger over album covers and to listen to different types of preview clips of songs to be used for creating a playlist (**P6**). The initial response by 62% of the participants was very positive. The prototype was described to offer an innovative and fast way for music discovery. The concept was said to support music discovery without having a deeper knowledge of musical terms and attributes, and getting instant feedback for user actions in the album cover space. The exploration by moving around in the album cover space and listening to typical preview songs for different styles received very positive feedback. The prototype started with an almost black screen and the user needed to place a finger to the screen to get started. This was confusing for some users but once they got started, the prototype was considered easy to use. The prototype was highly ranked in terms of being fun, looking impressive and presenting music in a novel way. The overall grade for the prototype was very good and even slightly higher in longer-term use. Still, the prototype was seen to lack detail for longer-term use and the users wanted to expand the possibilities in music discovery by enabling access to full albums and performing textual searches. Despite the lack of detail in the prototype, 63% of the users wanted to continue using it, mainly to complement other music applications (66%).

4.1.7 Mood pictures for music discovery

The first mood picture prototype used preset mood pictures and automatic playlist generation for discovering music (**P7**). The first impressions of the prototype were clearly positive for 63% of the users who stated the concept to be fun, novel, and a new way to explore music beyong genre boundaries. The users felt that their mood played an important role in music discovery. Navigating between the different preset mood pictures and listening to the associated music received good feedback from the majority of the users. Still, 37% of the users found the prototype to lack flexibility in terms of having to browse only between the selected mood pictures and the associated music. This was especially highlighted in longer-term use of the prototype. There was not enough user control nor matching mood pictures for all the users. Despite the restricted implementation, the concept was favored by 85% of the participants and was said to be well suited to exploring different types of music. All the users were unclear as to what mood the pictures represented and thus the music was not always seen to be matching the picture: this was partially caused by the mood picture design and partially due to the use of an automatic content-based music recommendation system for selecting the seed songs.

The second mood picture prototype, called Collaborative mood pictures (MoodPic) (**P10**), still offered preset mood pictures, but was newly designed with initial hand-associated songs. The essence of the new prototype was the ability for the users to add their own pictures as a basis for a collaborative mood playlist creation over genre boundaries. The preset pictures were used to guide the users and to provide fast access to different types of music. The concept of collaboratively populating mood picture playlists was seen as a great idea. The prototype service was seen to help in music discovery, offer a novel music listening experience, novel ways of sharing music, and to be well suited as a way to listen to music. The majority of the users were interested in organizing and searching music based on the mood pictures. Having others' contributions to complement a user's mood playlists was said to provide a very good experience and was found interesting and fun. The concept was well understood and fulfilled the expectations of most users. The majority of users were interested in using the service in the future and were willing to recommend it to their friends. Many considered collaboratively created playlists to well match the pictures and contain the types of music that the users expected it to contain. Most importantly, the users were willing to invest time and careful consideration in selecting song additions to others' playlists and uploading their own pictures as a basis for collaborative playlists. The users would have wished to have broader social interaction possibilities in the service, which affected the overall user experience. Additionally, some users were slightly worried about the lack of playlist content moderation and privacy with a potential large-scale commercial system.

4.1.8 Conquering physical territories for music discovery

The MyTerritory prototype enabled users to circulate and conquer physical territories from a map interface and associate music to these territories (**P9**). The user-defined territories could be of different shapes and sizes limited by the streets and the route of the conquering user. The prototype featured different ways to consume music associated with these territories: playing a territory, selecting an area from the map, and playing songs along a route on the map. The users appreciated the combination of outdoor activity, music discovery and gaming. The competition aspect in conquering territories and populating them with music split opinions almost evenly. Half of the participants preferred collaboratively associating songs to these territories without the competitive aspect. The overall concept was easily understood by most participants and it was found interesting and novel. Most users found it very interesting to see what types of music ruled each area and the service was seen to offer a novel way to share music. The prototype implementation was not perfect and did not work flawlessly, which affected the overall user experience. Still, most users could appreciate the concept over the flaws in the technical implementation and were interested in using the service for music discovery also in the future. Users wished to know more about the other players and their progress in the game. The current implementation only showed a nickname and the associated song for each territory, which was found to be restricting. The three proposed ways of selecting music from the map interface were generally seen favorably.

4.1.9 Summary of user experience in novel music discovery prototypes

The results show that the most positive user experience related comments were concerned with the novel and visual ways to discover and access music catalogs. The highly visual appearance and the strong use of stereotypes and visual metaphors were appreciated. Many positive comments also related to the simplicity of the prototypes and the ability to get quick access to different types of music. The main negative feedback was related to not supporting all music consumption habits and browsing styles such as textual search, especially in longer-term use.

Based on the individual prototype findings, the concepts MyTerritory, collaborative mood pictures and adaptive avatars presented in **P9-P11** offer a better overall user experience compared to others due to their wider support for different music discovery modes. Those three concepts offer basic music searching capabilities in addition to playful music discovery and therefore do not seem to include big hindrances regarding the support for different music consumption habits (see Section 2.4).

4.2 What kind of playful concepts work well for music discovery?

In Korhonen et al. (2009), the important characteristics in playfulness include pleasurable, rewarding or reinforcing experiences and that the behavior (in a system) is exaggerated. The key playful aspects for the researched prototypes are based on a personal analysis of the prototype evaluation results and observations during the user trials. The playful aspects with related PLEX model categories (Section 2.5.3, derived from Arrasvuori et al. 2011) are presented in Table 5.

Prototype / Publication	PLEX-categories	Key playful aspects
Virtual World (**P2**)	Captivation, Discovery, Exploration, Humor	• Potentially captivating experience in the virtual world city with the genre-based area visualizations • Humorous characters to be encountered on the streets including, e.g., cartoon characters • Music recommendations and discovery based on the stereotypical appearance of the characters
Avatar (**P3**)	Captivation, Discovery, Exploration, Humor	• Strong stereotypes in avatars and backgrounds to create a humorous and captivating exploration experience • Interactions for mixing avatar head, torso, and background for easy music exploration with the resulting playlist • Avatar and the related background create the overall theme of the prototype while consuming music for a humorous and captivating discovery experience
Potentiometers (**P4**)	Captivation, Discovery, Exploration, Humor	• The use of strong visual metaphors and humor in the theme based on selected genre • Experiences through a visual theme influencing the appearance of the potentiometer controls • Discovery by exploration using two large potentiometers
Cube (**P5**)	Captivation, Discovery, Exploration, Humor	• Mapping musical content on the faces of a cube for novel exploration of music and song related information • Playfull interaction of cube rotation for accessing and discovering music • Humorous mapping of cubes in the large area of miniature cubes (i.e., cube space) forming genre names with different colors
Album cover space (**P6**)	Captivation, Discovery, Exploration, Humor	• Moving a finger on the album cover space for exploring preview clips of different types of music • A halo spotlight and an animated waveform visualization for humorous uncovering of the album covers (and preview clips of songs) from the black background • Album cover space exploration and related interaction for captivating discovery experience
Mood pictures (**P7**)	Captivation, Discovery, Exploration, Humor	• Use of strong stereotypes in mood pictures for a humorous and captivating discovery, e.g., half-naked man in a sauna • Small animated elements in the mood pictures for enhancing the mood and thus the discovery experience • Ability to explore music by swiping between different mood

		pictures and consuming associated music instantly
MyTerritory (**P9**)	Captivation, Discovery, Exploration, Expression, Humor, Competition, Completion, Thrill	• Conquering (and re-conquering) physical territories with music for competition, expression, thrill and completion • The captivating experience of music discovery from the physical territories • The humorous way of consuming music from the territories with interactions common to navigation applications, e.g., routing • Collaborative playlist discovery from the territories for a new and exiting way for music discovery
Collaborative mood pictures (**P10**)	Captivation, Discovery, Exploration, Expression, Humor, Submission, Sympathy	• Sharing emotional feelings with user-generated mood playlists (Sympathy) • Collaboration in building mood playlists for discovery and exploration (also Submission and Expression) • Use of strong stereotypes and visual metaphors in preset mood pictures for easy recognition and humorous music discovery • A captivating music discovery experience using the associated mood picture as a background image for mood enhancement
Adaptive Avatar (**P11**)	Captivation, Discovery, Completion Exploration, Expression Humor, Thrill	• Populating the initially naked avatar character with music listening for completion and expression • A captivating and potentially thrilling experience to follow the constantly evolving avatar appearance • The use of strong stereotypes and visual metaphors in avatar appearance for a humorous visualization of the recent music listening history • Ability to explore other users' avatar characters and listen to the associated music • Sending visual recommendations (avatar parts with associated music) as a humorous social interaction and music discovery method

Table 5 Mapping of concepts to PLEX-categories (Arrasvuori et al. 2011) with playful key features.

All the playful prototypes in this thesis focus on at least four categories from the PLEX-model: captivation, discovery, exploration, and humor (Table 5).

Based on the individual prototype evaluations in the previous section, most of the different proposed approaches for music discovery had mainly positive first impressions.

The SuperMusic prototype (**P1**) performed very well overall but this was before, historically, streaming mobile services became common. Additionally, the prototype is not as playful and visual as the other approaches and thus ignored in this section. The virtual world approach (**P2**) was not well perceived and using a virtual world solely for music discovery is not encouraged. The virtual world approach was said to work best when combined with another main activity such as playing a game and having the music discovery as a side-product. The virtual world was said to offer many new opportunities for music discovery in **P2** but they were not taken advantage of in the implementation.

Prototypes **P3-P7** were compared in **P8**. The overall winner of the comparison was the album cover space (**P6**), which was considered an almost ready-made concept suitable for music discovery. The median value of the overall performance evaluation even increased slightly in longer-term use. Users appreciated the aesthetic look and simplicity of the playful music discovery by moving a finger on the album cover space, listening to preview clips of songs, and the ability to go beyond the initial playlist recommendation using a playlist song as a basis for a new playlist. In this comparison, the cube approach was found the least preferred option for music discovery mainly due to its interaction problems and the confusion in the operational flow of the prototype implementation. Still, users also saw potential in using cubes for music discovery as an overall concept, e.g., being able to see two types of information at the same time and being able to change these views. The greatest potential as a concept in the comparison study was identified with avatars (**P3**) and mood pictures (**P7**), where the implementation shortcomings were highly affecting the overall performance. With avatars (**P3**), the users wanted to have more control over the playlists, more variation in the avatar selection, more clarity in the relationship between the avatar visualization and the resultant songs in the playlist, and the possibility to make the avatars more personal. Mood pictures (**P7**) were seen as a very interesting approach for music categorization but the users wanted to customize the pictures and change the music associations to better match their preferences. This feedback was taken into account with the new iterated prototypes.

The new improved prototypes (**P10** and **P11**) were very well received and the use of mood pictures and evolving avatars can be highly recommended for music discovery. In **P10**, the ability for the users to upload their images as a basis for mood playlists and populating other playlists in the system with appropriate music were the highlights of the concept and showed deep user devotion and enthusiasm. In **P11**, being able to see the constantly evolving playful avatar character was seen as a great idea and it resulted in unexpected new ways of using a music discovery system in terms of different avatar creation strategies. The evolving visual appearance of the avatars was driving user behavior in terms of music consumption. Additionally, visualizing users' recent musical taste with a playful avatar character with quick access to the music was found especially nice.

The social music discovery by playful gaming in an outdoor environment was very much liked as a concept and should be researched further (**P9**), but the competition aspect in conquering territories split opinions and the approach should be carefully considered. Generally, the combination of outdoor activities and music discovery was seen to have potential.

The results show that visually interesting and playful music discovery solutions can inspire users and work well, especially for exploratory music discovery. Even very simple visual implementations can be found fun, entertaining and suitable for quick music discovery but the experience will not last without the support for both active and passive music discovery.

Based on the results, the most successful music discovery services, i.e. (Album cover space presented in **P6**, Collaborative mood pictures in **P10,** and Adaptive avatars in **P10**) include several key characteristics:

- **Simplicity in design**
  - Limited amount of features in the user interface.
- **Quick access to music**
  - Ability to start playback of different types of music quickly, e.g., moving a finger, playing an avatar, and playing a mood playlist.
- **Playful discovery experience**
  - Captivating and humorous approach for music discovery, e.g., finger movement in an album cover space, user-generated mood pictures and playlists, and adaptive avatar appearance.
- **Highly visual appearance**
  - Strong emphasis on visual appearance (**P6, P10, P11**) and the use of strong stereotypes in design (**P10, P11**).
- **Feedback for actions**
  - A responsive system for user actions, e.g., releasing a finger in album cover space to create a playlist, collaborative playlist creation, and changes in adaptive avatar appearance.
- **Support for collaboration, social interaction, and self-expression** (**P10, P11**)
  - Allowing rich interaction with other users via the system, e.g., collaborative mood playlists and listening to others' avatars.

The results indicate that well designed visual music discovery systems that support different music discovery modes (active, passive, and exploratory) have the potential to replace existing music discovery solutions as the main interface for new music. From the implemented prototypes, **P10** (Collaborative mood pictures) and **P11** (Adaptive avatars) are the most promising concepts to be used as a basis for a full-scale music service. Regarding **P10** and **P11**, the users were interested in using the prototypes in the future and wished to see the key elements in commercial music services. The users were not expressing any major hindrances that would risk the use of such concepts as the main interface for music (in contrary to **P2-P7**). Additionally with **P10** and **P11**, the users were willing to invest a significant amount of time to choose quality content for the system (in **P10**) and to steer the avatar development (in **P11**). These clearly indicate a genuine interest towards the concepts. All the other playful approaches (except for SuperMusic and Virtual worlds, **P1** and **P2**) work well to complement other music services, bringing a playful approach to music discovery.

4.3 What are the design implications for music service developers?

Based on the prototype evaluations, this thesis proposes a set of design implications for music service developers having a strong focus on music discovery. The proposed design implications work as a checklist for developers, from the early concept creation phase to later evaluation. The design implication topics include easy discovery, collaboration, social connections, sharing, creativity, personalization, and physical activites. Every included publication contributes to the design implications.

4.3.1 Encourage easy and entertaining music exploration

The results show that active music listeners can easily be motivated into playful music exploration and that the system should respond with easy-to-use but entertaining tools to support it.

Design implication 1: Simplify the design for discovery

Simple intuitive interactions, lightweight browsing, and the ability to quickly discover music without broad musical knowledge were highlighted in the results (**P2-P7**). Many ordinary listeners have a limited knowledge of musical terminology, e.g., genres. The results suggest having a simple entry-point for music discovery with strong emphasis on visualization to relay the correct message to ordinary listeners. Further adjustments, variation, and details can be offered for music enthusiasts. The simplicity in design in a positive way was the most highlighted in (**P6**). One of the key elements is quick access to music when needed (**P3-P7**). The simplicity and intuitiveness are supported by Merčun & Žumer (2010) in the field of information visualization. They conclude, based on their study on visualizing for exploration and discovery: "*The biggest mistake with information visualization in the past has probably been the fact that visualizations have been too complicated to intuitively understand and very often too extensive to enable effective manipulation and exploration.*" (Merčun & Žumer 2010, p. 7)

Design implication 2: Fade genre boundaries

Genres are currently the most used way of categorizing music but it does not mean that it is the only and best way for music discovery. The results suggest allowing music discovery beyond traditional genre boundaries to enable richer exploration. The results indicate a very good experience by using mood pictures for categorizing playlists (Collaborative mood pictures, **P10**). Using moods for music categorization can bring the emotional aspect of songs as the key attribute in music discovery. For example, there are songs from many different genres that can match angry or sad moods. Pictures representing moods can be interpreted differently and this may result in rich and versatile playlists as in **P10**. When using pictures to represent moods, the design should aim to focus on one specific emotion in order to help users to correctly interpret the basic meaning. If the picture is too general or can be interpreted in too many different ways, the resulting music might provide a bad experience for the user due to confusion (e.g., sauna picture in **P7**).

Design implication 3: Offer multiple ways to access content

In the studies, many different ways of music discovery were identified. Most of the researched prototypes were initially unclear and even confusing for a minority of the participants (**P2-P7**). This suggests providing multiple ways of accessing and discovering music and letting the user find the most suitable channels over time. Most users are very used to browsing music based on genres and thus complementing a system with new categorization approaches while still maintaining access with the traditional genre classification can be useful.

There are many different ways to listen to music, e.g., listening to a mix of top hits from different artists, random listening, listening music solely from one genre, searching for a specific song manually and listening to full albums from interesting artists. The system should support different music consumption habits to enable finding interesting music and listening to music once it has already been found. The three music discovery types (active, passive, and exploratory) should be addressed if the service focus is on being the only interface for music discovery. The different music discovery types are described in Section 2.4.

Design implication 4: Offer statistics, lists, and popularity information

During almost every study the interest towards seeing statistics, top-lists and what is popular at the moment came up in the interviews regardless of the used discovery method (**P1-P11**). The statistics and top-lists include both personal and other users' data. During the SuperMusic trial (**P1**), the emailed statistics and top-lists were highly appreciated by most users and are even more highlighted when there are otherwise insufficient means for music discovery. By providing user-related top- and activity-lists the service can encourage even more activity and give credit to users for their activity by giving visibility in the system.

Design implication 5: Use less words and more visual elements

Replacing unnecessary textual elements with entertaining and playful visualizations received very positive feedback from most users overall (**P2-P11**). A highly visual appearance in the prototypes was said to make the feel straightforward, simple, intuitive, and entertaining. The basic visualizations should be very simple and targeted to non-musical experts. Smaller details and visual cues can easily be added to enhance the basic visualization and to bring more depth and expressiveness while still maintaining easy identification for bigger user masses. The visual appearance should never be a blocker for quick music playback but a complement and an enhancement to the system. There should be the minimum amount of steps necessary to get good music playing (**P1-P11**).

The use of strong stereotypes in visualization helps in recognition and can offer a more entertaining and humorous approach to music discovery. This was shown particularly well in the use of avatar characters in **P3** and **P11**. Using human-like characters can also be a sensitive domain though and the use of strong stereotypes might offend some people, see Holm (2012, Section 7.4) for details.

Animation is a powerful tool to capture a user's focus and to deepen the meaning of visual content (**P7**). The small experiment in (**P7**) with animated content suggests to use them to enhance otherwise static images when trying to boost the emotional side of the service.

Design implication 6: Utilize playful aspects in music discovery

Including playful aspects in music discovery services can enhance the overall experience and include even humorous elements in music discovery (e.g., Adaptive avatars, **P11**). Music discovery should be designed as a fun task and the design should aim to provide enjoyment over efficiency. Music discovery does not need to be serious. Any music discovery concept should have the right balance between quick access to preferred music and playful elements. If the music discovery task is too time consuming and cumbersome due to playful elements, the overall experience will be bad (Virtual world, **P2**).

Captivation is a key element for a successful music discovery service and incorporating playful elements may enhance user engagement and thus captivation. Being a part of a larger structure or sharing emotions and feelings through music and pictures as in **P10** (Collaborative mood pictures) seem to increase the engagement. The use of evolving content (DI. 13) should be considered as a tool for playful experiences.

Competition and completion are elements of playful experiences that should be brought to music discovery systems as an optional addition. Forcing all users into competition might result in less engaging experiences for some users (MyTerritory, **P9**). Still, many also prefer competetive elements in music discovery services and they could provide novel ways to discover music.

The concepts introduced in this thesis utilized competition, completion, discovery, exploration, expression, humor, submission, and sympathy from the playful experiences framework of 22 categories (Table 3, Section 2.5.3). The use of these selected categories provide playful experiences for music discovery. Other categories need further exploration.

4.3.2 Connect users and support collaboration

Enabling social interaction and collaboration between users is important in music discovery systems.

Design implication 7: Connect users with music-centric social interaction

The social contacts between users were also seen to be very important in music discovery applications (all concepts, **P1-P11**). An important design implication regarding the use of a social network for music discovery is to have a music-centric focus to social interaction. There are already a number of popular social networking services and the users were interested in music focused interaction to avoid overlap with the other services. A highly focused system can still interact with, and contribute to, other social networking services.

Users were interested in being connected to others in music discovery applications. The connections between previously unknown users are encouraged so as to enhance the music discovery experiences. Music-related attributes such as musical taste, similar recently played songs, or similar personalized visualizations (Adaptive avatars, **P11**) can bring more meaning to the interaction and promote music discovery.

The enhanced interaction between the users and a strong sense of community and group dynamics may result in increased activity in the system and a more engaging music discovery experience (Collaborative mood pictues, **P10**).

Design implication 8: Support user collaboration for music discovery

In addition to music-focused social interaction between the users, collaboration in music discovery was seen beneficial. The collaboration aspect was especially highlighted in **P9** (MyTerritory) and **P10** (Collaborative mood pictures). Joint activities with other users may give the system a more personal touch and thus be more appealing. In **P10**, the users expressed investing a significant amount of effort in choosing songs for others' playlists and selecting a picture as a basis for collaborative mood playlists. The joint activities may result in high quality content and new ways of music discovery. The quality of the collaborative playlist was seen to be very good in **P10** and collaboration with other users was seen rewarding. When others consume the collaborative content (at least above a certain threshold) the creators should be acknowledged.

The other aspect in user collaboration is the competitive aspect between the users. This was investigated in **P9** with collaboratively populating physical territories with music. The game-like competition in music discovery can be seen as interesting, playful and entertaining but it can easily split opinions. The results suggest not using competition as the sole driver for collaborative activities.

Design implication 9: Build a bridge between the music listener and the creators

Being closer to the artists and bands was mentioned in several studies, especially in (MyTerritory, **P9**). The users would like to feel closer to the artists they like most. This can be achieved by letting the selected artists' actions or new releases influence the playlists in the system. The musical recommendations and preferences of favored artists can be valuable and used in combination to affect automatic music recommendations. The artists can receive feedback for their music from the users to be able to, e.g., fulfill their listeners needs in upcoming releases. The collaboration with artists does not only benefit the users of the system but also builds the artist's image in the eyes of consumers.

4.3.3 Lower the boundaries for content creation and sharing

User-generated content can be very valuable for a music discovery service. The content creation and sharing should be made fun, entertaining, and easy in order to maximize the potential activity.

Design implication 10: Support and acknowledge user-created content

Users can be motivated to be more active in terms of content creation and sharing when they are credited or acknowledged in the system for their actions. Encouraging users is a key element for music discovery systems where the user-generated content plays an important role (Collaborative mood pictures and adaptive avatars, **P10** & **P11**).

Creating and sharing new content (e.g., playlists, recommendations) should be straightforward, fast and easy so as to lower the boundaries for active interaction. In the studies, many users wished for even richer ways to interact but some wanted the fast and minimalistic approach (all concepts from the playful concept exploration phase, **P2-P7**). The results suggest having the extended content creation or sharing features optional to encourage a wide range of users to be active.

Following one's own content and contributions as well as others' reactions to it were said to be very interesting and important. An example includes following the development of the user's own mood playlist in **P10** (Collaborative mood pictures).

Design implication 11: Offer interoperability between services and devices

To be able to use music discovery applications and the created content with different devices and other services was seen to be in a key role for a wider success of a service (all concepts, **P1-P11**). Music is consumed with a wide range of devices in different locations and environments and well-designed services support most of these usage patterns. For example, the user might not always look at the screen when moving outdoors but could still want to use the system for music discovery.

Content interoperability can be important when performing difficult tasks or exporting content from a service to be used elsewhere. Examples include being able to modify playlists with a specific device (SuperMusic, **P1**), and exporting an avatar or a mood picture into a Spotify playlist and sharing it in social networking sites (Collaborative mood pictures and adaptive avatars, **P10** & **P11**).

4.3.4 Boost creativity, personalization, and self-expression

Supporting creativity, personalization, and self-expression are significant in exploratory music discovery.

Design implication 12: Enable user control in music discovery

Based on the study results, the users appreciated having an increased amount of control for fine-tuning and adjusting the music discovery parameters. For example, adjusting tempo and gain settings within a selected genre using potentiometers (Potentiometers, **P4**) was found to be very good. For simplicity, not all settings need to be visible at the start but should be available for tailoring the result to match user preferences. The users highlighted the ability to give feedback to the system, especially to improve and to be in control of the recommendations (SuperMusic and concepts from the playful concept exploration phase, **P1-P7**). Instant reaction to the user's feedback was seen as important. Good usability and interaction design strongly contribute to the feeling of being in control of the system.

Supporting learning and curiosity requires finding the right balance between the familiar and the unexpected to create an appealing overall system.

Another aspect in user control relates to the visibility and moderation of user-generated content and how it is used in the system. Users felt that they want to be aware of what is exposed from them to others, especially concerning user-generated content. The amount of non-adjustable exposure of user content may correlate to the user activity in creating content in a system or to how the content is created. An example includes fully open collaborative playlist creation based on users' mood pictures (**P10**) in relation to what type of pictures are chosen. Another good example is found from **P11** regarding adaptive avatars representing users' recent music listening and how the users want to appear in the system to others.

Design implication 13: Offer evolving content

When users are discovering music, it can be highly rewarding to see the visual side of the system evolve over time as they discover and consume music (Album cover space and adaptive avatars, **P6** and **P11**). The evolving visual side can enable the users to quickly and intuitively see the result of their actions. Based on user feedback, the system feels more alive and captivating when there is a constant change based on user activity or music consumption. The evolving aspects in a system also encourage users to frequently check and re-visit the system to be up-to-date, as with collaboratively created playlists in (**P10**).

The ability to iterate recommendation results in music discovery was seen as one type of evolving content. The option to take a song from the recommendation playlist and use it as a seed for a new playlist was well received (e.g., Potentiometers **P4** and Album cover space **P6**). This enabled users to go beyond the initial recommendations and to develop the playlist based on their actions.

Design implication 14: Provide means for personalization and self-expression

With visual music discovery applications, supporting personalization and self-expression has a significant role. Creativity and self-expression can be boosted by enabling different ways of using the system, being able to customize the system functionality, its appearance, and what is exposed to others (Collaborative mood pictures and adaptive avatars, **P10** and **P11**). When given the right tools, users will find innovative and unexpected ways to interact with the content. Supporting creativity and self-expression in the right way may result in increased creation of valuable user-generated content in the system (**P10**).

A visual music discovery service should somehow represent the user, which can be enabled by letting users give their own "touch" to the service. This can include modifying the service appearance or adding personal content to the users' own page.

Users saw it important to be able to show their musical identity to others and this should be supported in a music discovery service. The user's musical identity in a music discovery system may include, e.g., a music-related profile, listening history, preferred music, user's generated content, and music-related activity in a system. Musical identity can be promoted by means of, e.g., exposing user's musical taste and progress, identifying and visualizing a user based on the consumed music, and enabling rich music-related user interaction in the system. The users' musical taste may change over time and mood changes can be even more frequent. The system should live with the user and support and reflect these changes.

Design implication 15: Quality before quantity to enable a rich discovery experience

A limited community size and a possibly limited number of songs in a music discovery service can be compensated for with high quality content (i.e., relevant music for the users with enough variation) that is presented and accessed with creative and innovative ways (Collaborative mood pictures and adaptive avatars, **P10** and **P11**). The users have limited capabilities for consuming music and the focus should rather be to provide a pleasant high quality discovery experience instead of providing possibilities to browse all the music in the world.

Design implication 16: Involve physical activities in music discovery

The last design implication emphasizes the positive feedback related to the combination of music discovery and physical outdoor activity (MyTerritory, **P9**). Creating new music discovery experiences in either well-known or new surroundings can enhance the activity. The physical aspect in music discovery may bring new depth to already known surroundings and introduce a new positive aspect to daily routines (e.g., by conquering and exploring musical territories, **P9**). Incorporating physical activities in new surroundings can combine the music discovery with physical exploration. The physical exercise should not be the key factor in music discovery when reaching larger masses. Alternative ways for discovery should also be supported.

4.3.5 Support for the design implications

There is a very limited number of studies of visual music discovery concepts and prototypes where a clear list of design implications would have been proposed.

Weigl & Guastavino (2011) summarize user studies in the music information retrieval literature highlighting several key design implications for MIR system design from the themes of: 1) Undirected browsing, 2) Goal-directed search, 3) Recommendations on metadata, 4) Social aspects, 5) Organization of music information, 6) User interface appearance, 7) User support, and 8) Hardware/portable MIR device. The implications are well in line and show strong support with, e.g., DI 1, 2, 5, 7, 8, and 12 proposed in this thesis.

A study by Laplante & Downie (2006) concerning everyday music information seeking among young adults proposes several design implications including the following themes "*Informal channels/Distrust of experts*" and "*Music information seeking as a non-goal oriented activity/browsing over searching*". These themes and the related implications are well in line with the DI. 3, 7, and 8 proposed in this thesis. The implications in Laplante & Downie (2006) are broad and the themes can be seen to be extended in DI. 1, 5, and 11. In Cunningham et al. (2007) regarding finding new music, several design implications for MIR system design are proposed. The implications highlight, e.g., serendipitous browsing, mobility, including related information, unifying different forms of discovery, and social interaction. These implications are well in line with DI. 4, 11, and 12. In an earlier study by Cunningham et al. (2003), observed behavior in physical music searching and browsing included for example the following themes: "*Searching and browsing are interleaved* , "*Known-item searches are conducted*", "*Browsing is a significant activity*", "*Shopping is often collaborative*", and "*Music shopping is surprisingly visual".* These themes have a good match to the proposed design implications, e.g., visual aspects, collaborativeness, browsing being a significant activity to be supported, and exploration by listening clips of music for verification and discovery.

Hekkert (2006) has explored principles of pleasure in design and proposes four general principles of aesthetic pleasure: *"1) maximum effect for minimum means, 2) unity in variety, 3) most advanced, yet acceptable, and 4) optimal match."* The principles have a good match for the proposed design implications in this thesis. Some have a clear link, e.g., maximum effect for minimum means to DI. 1. Some links need a small amount of interpretation, e.g., 'most advanced, yet acceptable' refers to people being attracted by new, unfamiliar, and novel solutions while preserving the typicality, matching at least the DI. 3, 5, 6, 7, 8, 11, 12, 13, and 14. Another example is 'optimal match' encouraging to explore the use of different senses, e.g., for enjoyment and inspiration. The principle has a relatively good match to DI. 16, encouraging physical activities in music discovery.

Yu & Nam (2014) have recently studied the creation of humorous products proposing nine design principles including, e.g., "*Destructive play*", "*Visualization of taboo*", "*Zoomorphism*", and "*Abused product*". Many of the implications involve breaking barriers, providing unexpected outcomes, and breaking social norms.The work and principles are in line with the approach with music discovery prototypes introduced in this thesis. The proposed nine design principles support, e.g., DI. 6 and 14 in this thesis.

4.3.6 Summary of design implications

This section presented 16 design implications for music discovery service developers. The proposed design implications are music discovery specific and cover a broad range of design aspects to aid the design and validation of such systems.

The design implications are summarized in Table 6 with short descriptions.

#	Design implication	Short description
Encourage easy and entertaining music exploration		
1	Simplify the design for discovery	Support simple and easy discovery to reach a broad range of users
2	Fade genre boundaries	Enable new ways for music categorization
3	Offer multiple ways to access content	Support different music discovery and consumption habits
4	Offer statistics, lists, and popularity information	Engage users for music discovery by providing information about the trending and popular songs/artists and statistics related to the user's actions
5	Use less words and more visual elements	The use of strong stereotypes and visual metaphors can engage the user
6	Utilize playful aspects in music discovery	Avoid too serious music discovery and consider the use of humor and other playful aspects for captivating discovery experiences
Connect users and support collaboration		
7	Connect users with music-centric social interaction	Do not try to replace existing social networking services but focus to music-centric social interaction
8	Support user collaboration for music discovery	Provide richer content and increase engagement through collaborative activities and collaboratively created content
9	Build a bridge between the music listener and the creators	Bringing artists visible in music discovery systems benefits both parties
Lower the boundaries for content creation and sharing		
10	Support and acknowledge user-created content	Provide means to get unique and high quality user-generated content for differentiation and providing great discovery experiences
11	Offer interoperability between services and devices	Allow users to mix and match content between services and use different devices for discovery
Boost creativity, personalization, and self-expression		
12	Enable user control in music discovery	Let the user fine-tune the discovery experience for best results

13	Offer evolving content	Engage users with evolving content and provide visual feedback for actions to enhance the user experience.
14	Provide means for personalization and self-expression	Enable personalization and channels for self-expression and creativity to engage the user and inspire others for music discovery
15	Quality before quantity to enable rich discovery experience	Compensate the lack of millions of users or songs with high quality content
16	Involve physical activities in music discovery	Support physical activities for more contextual and novel music discovery experiences beyond traditional services

Table 6 A short summary of the proposed design implications for music discovery systems.

5. Discussion

This chapter first addresses the reliability and validity of the results and proceeds to describe the sensitive nature of music and its potential effect to music discovery systems. Concrete proposals to further enhance the design of the created prototypes are given for an even better music discovery experience. Finally, the chapter concludes with new approaches and future opportunities aimed towards an “ultimate” music discovery experience.

5.1 Reliability and validity of the results

Reliability refers to being able to obtain the same results on repeated trials with regards to an experiment whereas validity, in the area of design research and experimentation, concerns the study of being able to answer the right questions (Nielsen 1994). Validity can be divided into internal and external validity. Internal validity in experimental research concerns aspects related to the study measuring and addressing the right issues (e.g., research questions, methods, bias). External validity relate to generalizability of the results across different situations and/or people (Cambell & Stanley 1966, Isaac & Michael 1971).

Typical real-world examples of problems in measuring user experience qualitatively include: having the wrong type of participants in a user trial, performing the wrong tasks, and not taking into account all the relevant factors regarding the research topic. Statistical measures can be used to measure reliability with quantitative data. Still, the applied research methods need to be understood correctly and used accordingly to obtain valid results.

The participants in the user evaluation studies were mostly 20-40 years old. The users in the performed trials were selected based on their music listening activity, which was seen to have the biggest impact in measuring user experience and gaining valuable future development directions with visual music discovery prototypes. After all, in order to judge and evaluate different music discovery solutions properly, the users need to explore and listen to music with them. As the minimum requirement for user selection was active music listening with different devices and services, it was easy to see how the proposed solutions reflect the existing commercial solutions. The users who seldom listen to music and do not use online music services actively were not considered as the main target audience.

The studies did not include strict restrictions for age and the calls for participation were relatively open. The main channels for recruiting people were bulletin boards at Tampere universities, social networking services, music-related mailing lists, company-internal mailing lists, and personal recruitment. Most users were young adults and quite technology oriented. There were not many children and no elderly people volunteering for the trials and therefore the results are mainly applicable to young adults. The results from the small number of underaged participants (12-17 years) were generally in line with the other results, as were the results from the older participants (50+ years). Many of the general design principles most likely apply to larger groups of people regardless of age. The selected users well represent the biggest, by far, user group in streaming online music services, e.g., Spotify, where young adults (18-24 years) dominate the demographic distribution in the UK (Spotify Introduction 2011). All participants were Finnish and the studies did not address any cultural aspects of the music discovery prototypes. In addition to the users, the design focused on the use of Western stereotypes and colors. Therefore, the results apply mainly to active music listeners from high-tech oriented Western population.

All the trials included a period of free-use with the prototype and the users were able to explore the prototypes as they would outside of a user study situation. Some of the studies include comparisons between initial experiences and longer-term feedback (**P3-P7**), while others included a short usage period followed by the interviews. Only the virtual world concept (**P2)** was used during the interview sessions only.

The initial feedback for prototypes studied in **P2-P7** (playful concept exploration phase) was gathered during one long interview session. While the order of the tested prototypes was randomized, the ability to try many competing alternatives during one session may have had an effect on the user comments and experiences. Being able to try multiple different solutions at once gave an opportunity and a baseline for thinking outside the scope of only one prototype. The impact is seen in the plurality of future development ideas for each approach, partially influenced by the experiences (or lack of them) regarding the other prototypes. The users in this trial were then able to continue using prototypes (**P3-P7**) freely for five weeks. Quantitative user feedback was gathered using a seven-point Likert scale. Based on the correlation of results between the different prototypes for initial experiences and longer-term use, the quantitative results are valid and reliable. Nine additional participants filled an AttrakDiff questionnaire for the virtual world prototype (see **P2**, Chapter 4). AttrakDiff is a 28-step questionnaire for evaluating the pragmatic and hedonic quality of products using a seven-point semantic differential scale. The results are well in line with the other participants.

The quantitative data for the prototypes in playful concept exploration and iterative concept design phases was gathered using a seven-point Likert scale and the results were visualized using boxplots (**P2-P11**). Statistical analysis was performed in order to find out whether the answers depend of, e.g., gender or age. Chi-Square was initially used to investigate the problem. If the conditions for using the Chi-Square test were not met, Spearman's correlation and Wilcoxon's signed rank tests were calculated. The requirements for using Chi-Square include that no more than 20% of the counts (i.e., response options) have less than five answers and all counts have at least one. This condition was not commonly met with the answers to our UX questionnaires. Instead, Spearman's correlation was calculated address the issue. Wilcoxon's signed-rand test was used to investigate the differences between the answers initially and after longer-term usage.

Some of the prototypes were used for a shorter time period, e.g., **P9** (MyTerritory) for one session, **P10** (Collaborative mood pictures) for 2-5 days, and **P11** (Adaptive avatar) for 1-3 weeks. The shorter usage time may have had an impact on the results, especially for **P9**. The rapidly evolving content, bigger amount of simultaneous users and a limited set of features in **P10** and **P11** enabled the users to experience the systems in detail and repeat similar tasks many times during the trial period. The amount of user-created playlists and associated songs in **P10** and listened songs in **P11** support the validity of the results in terms of having enough experience to evaluate the systems.

There was no control group in the studies (excluding recommendation evaluation in **P1**) for result comparison, i.e., a group that would use a basic text-based player and answer the same questions. This is a clear lack in the evaluations. Due to the vast number of field trials, the resources were focused into having a maximum number of participants for qualitative and quantitative feedback rather than dividing the participants into two groups (where the other group would not use the proposed music discovery prototype). Based on the interview feedback the users were well aware of different alternatives for music discovery and were capable of judging the proposed prototypes well and semi-professionally e.g. comparing the features and appearance to other existing solutions and thus delivering meaningful results. Concerning demand characteristics, the users were assessing the prototypes based on their music consumption needs and ways to discover new music. Many of the features and prototype characteristics faced a vast number of different responses from the users and many of them felt highly personal outruling many of the typical cases of demand characteristics. Additionally, many of the trials included a phase where the users were able to test the prototypes at home for a longer time alongside to other available commercial music applications. The field trial period was addressed carefully in the interviews including the experiences and potential problems with the proposed music discovery concept while discovering music in a natural setting.

Prototypes in **P1**, **P9**, **P10** and **P11** introduced social interaction between the users. The lack of a real-life social network in the services during the trial was seen as a potential risk while conducting the studies. Users reported being able to explore the social features with the prototypes with the other trial users and being able to get a good overall picture of how they would use them with their existing connections. Still, if the participants' real personal networks had been present in the services, the interaction could have been even livelier, allowing for further insights into social collaboration in music discovery.

Furthermore, the detailed design of the user interfaces plays a significant role in how the user perceives a music discovery prototype. Small changes to the user interface may have a big impact on user acceptance and preference towards the proposed music discovery approach. Some great technical solutions enabling accurate music recommendations might not be appreciated if the user interface fails to show the best of it in a usable and entertaining manner. For example, interaction problems with the cube in **P5** resulted in a bad overall user experience for some of the users. Additionally, the prototypes from the iterative concept development phase (**P9-P11**) included a very modest amount of songs (~500 songs). Although some of the user feedback addresses the use of the concepts with larger amounts of content and users, the results are valid only for small music databases. The concepts should be further evaluated with larger amounts of content. The large-scale use potentially poses new requirements for searching and filtering the content in the service.

The proposed design implications (Section 4.3) are focused on music discovery services. Many of the themes can be interpreted more generally and considered for enhancing the design of interactive discovery services in other domains. Examples of suitable themes include simplifying the design, multiple ways for accessing content, emphasizing the use of visual elements in the user interface, enabling user-control, and providing means for personalization. Some of the themes can help the designer to think in a broader domain, e.g., how to go beyond traditional categorization of books (*DI. 2*) or how to involve physical activities to photograph discovery (*DI. 16*).

Based on the presented details, the results of the thesis are valid for the cultural context where they were studied and for most potential user groups of music discovery services.

5.2 Music is sensitive and personal

It is not only music. Music can play a big role in emotions and past experiences. The psychological perspectives in (Juslin & Sloboda 2001) divide the primary sources of emotion in music to intrinsic and extrinsic emotions. In general, intrinsic emotions come from the inside of an individual, from the pleasure of performing a task whereas extrinsic emotions come from external rewards and sanctions. Intrinsic emotions in music refer to the relationship between the structural characteristics of music and personal emotional experiences in music at a specific point of time. Extrinsic emotions in music refer to linkage between musical structure and prior experience with emotion and linkage with experienced music and non-musical factors carrying their own emotional messages (Juslin & Sloboda, 2001).

Music discovery systems may be designed to offer entertaining and playful experiences, however, it is currently impossible to reflect the users' past experiences and values in a music discovery system automatically. These experiences and values can play a big role in music preference. For example, negative experiences in life can be associated with an artist or a specific song, which the user does not want to hear in any situation regardless of it being the best recommendation based on other listened music. On the other hand, positive experiences can be reinforced by listening to music associated to positive events.

The culture, religion, and musical identity related aspects may also have a big role in preferred music and thus in music discovery (Serra 2011, Tzanetakis 2007). Examples of unfortunate choices of music include selecting gospel for atheists or playing music from a musical talent competition winner to an ambitious member of a “cellar band” trying to succeed in his musical career.

Music listeners seem to fall into two categories in terms of the focus within a song: the music and the lyrics. This also plays a significant role in music discovery and has not been used widely in music discovery systems, partially due to the complexity of semantic text analysis (Goddard 2011). Semantic lyric analysis could help, e.g., in interpreting the mood of a song or in searching for a song about a specific topic.

5.3 Towards an ultimate music discovery experience

This section firstly discusses the potential directions for enhancing the existing prototypes towards even better music discovery experieces, mostly utilizing the future development ideas and observations from the prototype studies, complemented by selected personal design ideas. This is followed by a more general and future-facing discussion of taking the music discovery experiences to the next level.

5.3.1 Enhancing the existing prototypes

Depending on the usage scenario, the prototypes should support different ways for music discovery (Section 2.4). If the prototypes are used to complement other music players as exploratory music discovery interfaces, they can be highly focused and neglect traditional music player features. To use the prototypes as a primary interface to music, active and passive music discovery modes should be supported (see Section 2.4) including having the possibility to, e.g., listen a mix of top hits from different artists, listen in a random order, listen to music solely from one genre, search for a specific song manually and listen to full albums from interesting artists.

There is always the risk with visual music discovery concepts that users get bored with the solution or the concepts pose design challenges in the mobile domain with smaller displays. These factors should be taken into account in the design by providing enough depth and evolving content in the concepts without sacrificing simplicity. Potential future directions for enhancing the music discovery experiences using the current concepts are introduced next with a mapping to the related design implications in section 4.3.

Virtual worlds (P2)

Despite the users not preferring the current implementation, many saw the opportunities that a virtual environment could bring to music discovery. The whole application should be re-designed in terms of interaction, characters, and surroundings. The development should aim to provide more interactive content with the ability to interact with literally anything in the world. The enhanced version could include, e.g.;

- fighting with unpreferred artist representations to indicate musical preference (*DI. 6, 12, 14*)
- destroying areas of unwanted music appearing in the world (*DI. 6, 12, 14*)
- quick navigation by driving around the city while listening to car radio for previewing clips of surrounding music (*DI. 1, 3*)
- exploring numerous hidden features that are not obvious instantly (*DI. 13*)
- enabling different types of mini-games inside the world to keep the user engaged (*DI. 6, 14*)
- ability to personalize (and view) the user's own character in the virtual world with the appearance having an impact on the encountered people and surroundings (*DI. 14*)

The music discovery experience could have a clear goal and an interactive story to back it up. With rich storytelling, the user could feel that the actions play a part in a "game" and music could be discovered on the way (*DI. 6, 13*).

The virtual world should be rich from a sound perspective as well. The world could include ambient sounds, background music, and noises to aid exploration. Clubs and other places of interest could be heard from a distance and the soundscape could help the user to navigate (*DI. 1*).

A slightly different approach would be to navigate the virtual world using an aerial view with strong visualizations of different districts and interactive content. The user could hover to the preferred places quickly and zoom in to explore more details. The zoom level could be mapped to genre (or other high-level classification), artist, album, or song –level information to be visible in the world (*DI. 1*).

Avatars (P3 and P11)

In general, avatars work well for music discovery and have received a positive reception in both **P3** and **P11**. In is possible to design avatars that are recognizable by most users, which was also shown in the more detailed visualization study by Holm et al. (2010) using an extended set of avatars from **P3**.

One way forward is to combine the best of both approaches (**P3** and **P11**). The ability to create cross-genre playlists as in **P3** should be brought to **P11** but clearly showing the linkage between the songs and the avatar parts in the resulting playlist. When using avatars to represent music listening history, there could be more than one avatar to be able to see the recent listening history and the complete listening history (*DI. 3*). Additional avatars could be used to provide visual statistics and top-lists regarding the popular, trending, new, and recommended content (*DI. 4*). Users were interested in filtering out selected songs from appearing on the avatar and this should be supported. The fear of having all listened music mapped to the avatar may have a negative effect to some users of the system (*DI. 12*).

The service could include avatars representing and promoting well-known artists such as Madonna or Mick Jagger. These characters could offer music from these artists and/or music that these artists currently like to listen to (*DI. 9*). The well-known artists' avatars could evolve visually like others in **P11** and match the current state of the artist, e.g., a ghost of Michael Jackson could still be promoting music (*DI. 13*).

It would be interesting to experiment with less cartoon-like avatar characters, which were found slightly childish by some of the users, and to find a good middle-ground while still maintaining the strong stereotypes in the appearance. Alternatively, there could be several visual styles for the avatars, customized to selected target groups and cultures (*DI. 1*). Avatars and their relationship to racism is discussed in Holm (2012) and should be considerd in the design. It would also be an interesting experiment to use a set of avatars for visualizing different styles of music for children, taking strong stereotypes from the iconic characters from the childrens' perspective.

One concrete way to promote avatar usage is to offer interoperability with other music services. Any playlist from a music service could be visualized using the same avatar creation techniques and shared in, e.g., social networking services. Shared visual playlists might be more interesting to receive and view in social networking services than the current music postings from various services. Linking music with the avatar parts could be provided in different formats, e.g., as Spotify links to enable a larger compatibility between services. The only thing needed is an API that is able to generate avatars with different attributes from a list of songs. The only technical challenge is matching the metadata between different services (*DI. 11*).

Potentiometers (P4)

Potentiometers can be used well to fine-tune details for a playlist or song selection. The current prototype (**P4**) included an unnecessary step from the potentiometer tweaking to playlist generation. The potentiometer tweaking should automatically create the playlist for the user with no extra steps, as many of the users expected while using the prototype (*DI. 1*).

The scale of the potentiometers could be varied depending on the available choices for the current setting, i.e., the potentiometer rotation could change songs more rapidly or slower depending on the amount of suitable content (*DI. 12*).

One clear development topic is to seek new metadata that can be used to bring a wider customizable selection of potentiometers to complement the current selection. Instrument recognition is a very interesting algorithmic challenge (Fuhrmann 2012, Giannoulis & Klapuri 2013) and recent successes with the development of such algorithms could enable their use in such applications. For example, the user could be able to turn different instruments on/off and then fine-tune the music selection with potentiometers. The song could be visualized with pictures of the instruments, or even with a band playing the instruments for intuitive interaction with the music (*DI. 14*).

The visual themes for the musical styles were seen to work well and aid the user in understanding the scope for fine-tuning the playlist content. The use of visual themes could be explored further to include evolving elements based on user actions. For example, the theme for pop music could change gradually based on the listened pop music. Each theme could evolve separately from others and show richer information about the user preferences within a genre or style (*DI. 6, 13*). The potentiometers and the related dashboard could be available for other users for experimentation. This could bring a new powerful social aspect when used like adaptive avatars in **P11** (*DI. 7*).

Cubes (P5)

The cube as an interface for music discovery could be explored further and extended. An important driver for further extension is to first fix the cube interaction to be smooth and easy for the users (*DI. 1*).

The actual cubes could be visualized differently to allow easier recognition and better match the associated music. The cubes could have different themes and colors relating to the associated content. Richer interaction with the cubes could be experimented with such as combining multiple cubes as one or dragging and dropping individual elements from one cube to another (*DI. 1, 3*).

The genre map of miniature cubes (**P5**, Figure 3) does not differentiate between cubes of the same genre and the user needs to remember the cube locations to retrieve a specific cube. This view should be greatly improved to allow turning the miniature cubes and applying the previously suggested cube themes to the genre map view as well. The user could potentially zoom into the genre map to view more details for the cubes (*DI. 1*). Social aspects could easily be brought to an extended genre map using cubes from many users. The social interaction could work using the cubes as a medium and the owner (or the representation of the owner) could be visible on one face of the cube. The owner could work as a link to other cubes in the system (*DI. 7*). Popular, trending, and recommended cubes could be included in the cube map (*DI. 4*). Additionally, users could build their own cubes from scratch by associating songs from the music database to a cube (*DI. 10*). Each user could collect a set of preferred ready-made or custom-made cubes to their own area in the genre map and the selection of the cubes and/or related songs could be collaboratively extended, similar to the way playlists are extended in **P10** (*DI. 8*, *14*).

Album cover space (P6)

The album cover space –approach was found best by many in the comparison study of five music discovery prototypes (**P8**). The interaction of moving a finger to explore music should not be changed; however, the exploration possibilities could be extended.

The playlist generation process could potentially be replaced while enhancing the album cover space interaction. For example, the user could move their finger and release it to continue playing music of that type without the extra step in playlist generation (*DI. 1*). While listening to good music, the album cover space could show related information and offer possibilities to steer the listening experience towards, e.g., listening to the full album, finding more from the same artist, or finding similar artists to listen to (*DI. 3*).

The album cover space could be dynamic and extend towards different directions while the finger is moved. Different music regions could have an additional visualization in addition to the album covers to be easily recognized and the region size could adapt to the user’s listening preferences, e.g., most commonly listened to music could take a larger portion from the album cover space (*DI. 13*). The region content could be influenced by the music that other people consume within the system (*DI. 7*). The regions could easily step outside the genre categorization and dynamically respond to the user’s music consumption by building personalized regions (*DI. 2, 14*). When the user encounters non-preferred music while exploring or listening to a playlist, the user should have the ability to give feedback to steer the album space content in the future (*DI. 14*).

Mood pictures (P7 and P10)

The mood pictures are a good way of categorizing music beyond traditional genre boundaries. The way of exploring preset mood pictures in **P7** could be enhanced to allow customization of mood pictures and music associations while maintaining the existing interaction (*DI. 14*). The preset mood picture set in **P10** should be further iterated and the selection could be extended (*DI. 1*, *3*). The ways in which the mood picture can evolve based on the music consumption of the related playlist should be investigated. One alternative is to include an additional semi-transparent layer on top of the mood picture for evolving content such as collaborative contributions and dynamically changing animated elements such as rain and fog (*DI. 13*). Contextual attributes could be included in the evolution of the mood pictures thus bringing in more real-time elements and supporting physical activities (e.g., running, cycling) (*DI. 16*). While the user is engaging in an activity, the playlist selection could be automatically filtered to show the best matching mood pictures first (*DI. 1*).

While the amount of user-generated content increases, the focus shifts towards finding the relevant mood playlists from the system. There could be easy and visual ways for users to annotate the playlists regarding the moods they represent. This would allow filtering the playlists in the system. Alternatively, automatic mood analysis based on the playlists' content could be performed (*DI. 1, 5*). The most popular, trending, and recommended playlists should be found under their own categories for easy exploration (*DI. 4*).

Better playlist management options and different privacy levels should be introduced to the playlists to enable broader use of the system. While most playlists could remain open for additions by anyone, some could be more private, to be used within a closed group or by creators only. In addition to the playlist visibility, the management options for song contributions should be explored. One alternative is to have the playlists open, as implemented, or to be able to moderate the song additions before they become visible in the system for all users (*DI. 12, 14*).

Physical outdoor music discovery (P9)

The physical outdoor music discovery (**P9**) should incorporate both, competition and collaborative aspects (*DI. 6*). Building the playlists for different "territories" should acknowledge the ways the content was created (*DI. 10*). It was seen as a good way to motivate users into physical exercise. This should be supported (and even further promoted) in the future, but not as an obligatory element of the music discovery service. The territory conquering as a gaming element could be extended to include not only circulating areas but taking also altitude information into account. With the proposed extension, users could conquer more precisely, e.g., floors or sections from buildings. Having an increased amount of associated tracks for each building would create larger playlists for the music consumption side of the prototype (*DI. 14, 16*).

The control and management options for content in each territory should be extended to allow richer interaction and bring social aspects to be more visible. In general, users should be more visible in the system for others and not just music (*DI. 7*). The management options for musical content could also include the ability to mix songs between territories and view location-specific playlists easily and visually without having to access them separately (*DI. 1, 3*).

More contextual features could be easily brought in to extend the current concept. Users' playlists with other music player applications could be influenced by the songs attached to different territories allowing cross-application interoperability. The system could provide territorial recommendations using an open API to any existing music player (*DI. 11, 13*).

Local artists and record labels could be incorporated into the territory-concept and different entities, e.g., individuals, well-known artists, record labels could be visualized differently (*DI. 9*). Some songs or territories could stand out from the map view using 3D elements in the user interface. For example, territories with music matching the user's listening history or most listened territories could be taller on the screen like skyscrapers compared to ordinary office buildings (*DI. 1, 3, 5*). Satellite images could be used to make the territories more life-like and merge with the real world. In some cases, the music discovery concept could be viewed as an additional layer in a navigation and map software (*DI. 11*).

5.3.2 Further potential approaches towards the ultimate music discovery experience

The previous section focused on enhancing the proposed music discovery concepts for an even better music discovery experience. In the studies, there were always some individuals that did not like the proposed music discovery approach. An example is found from (**P8**, Figure 24) where the user responses for the best and worst prototype were scattered between the five proposed systems. The possibility of having multiple alternatives for music discovery should be investigated. The user could potentially personalize the music service fully to their music consumption and exploration needs. If the visual music discovery alternatives could be used with a common input, the user could switch between views at any point and use the most suitable browsing and exploration method depending on the use case. For example, the user could start by searching a song with its name and start playback. Then, the user could switch to a more exploratory mode and see related mood pictures or use the potentiometers to fine-tune the music selection using the initial song as a seed. At any given point, the user could switch back to the traditional textual mode. The system could be able to remember the user's preferences for interacting with different content and automatically adapt to the user's common usage patterns.

As mentioned in the previous section, richer metadata opens new opportunities for visual music discovery. Automatic instrument recognition, as an example, is a promising technology and is able to bring a new point of view to music discovery. Combining audio similarity and instrument recognition would work in use cases such as, e.g., “I want similar music to this but with electric guitar” or “I want something like this but without the banjo”. The opportunities for visualizing such use cases are almost endless. To go a step further, if the processing could address the amount and role of the instrument in a song the user could request for a track with more dominance of a specific instrument. This type of audio analysis could become reality within a few years. Research is needed regarding how to bring these new opportunities to the user in an entertaining and fun way without the need for a broad musical knowledge from the user.

Combining multiple modalities is a potential future direction to enhance music discovery experiences. The discovery interaction could be taken out of the display to an external object. The ability to turn, press, twist, smell, and even taste an object open entirely new music discovery opportunities. Different musical attributes could be mapped to the objects allowing simultaneous modification of multiple variables. The rapid development of 3D printing may speed-up the use of personalized physical objects in various application domains.

To go well beyond the current state of the technology, the music discovery should include a wide range of contextual, emotional, social, and technical aspects (many aspects that are being researched individually at the moment) brought into one easily approachable solution for the “ultimate experience”. For example, the user could be able to save a state where a song was experienced including all the attributes affecting the overall experience. A number of these states could be captured and later replicated precisely in a digital domain, and expanded to different directions. The user could further enhance the real-life experiences with music and re-live the precious moments again with, e.g., different weather, a different version of the song playing, different mood, or simply, ‘with a twist’.

6. Conclusions and future work

In this thesis work, novel concepts for visual music discovery were studied. The goal was to research the user experience (UX) of new types of music discovery services and find out key design implications to support the service development for music discovery. The research of music discovery concepts was divided into three phases: initial concept design phase, playful concept exploration phase, and iterative concept design phase. The thesis introduces, in total, ten different concepts for visual music discovery with implemented prototypes and user evaluations with Finnish active music listeners gathering both qualitative and quantitative data. This thesis contributes to both academic research and commercial music discovery service development focusing to Western music markets.

The results describe the user experience with different types of music discovery services, showing that the most positive user experience related comments concern the playful and exciting ways to discover and access music catalogs. The use of a highly visual appearance, visual metaphors and strong stereotypes in the prototype design were appreciated and encouraged. Many positive comments also related to the simplicity of the prototypes and the ability to get quick access to different types of music without a broad musical knowledge. The main negative feedback was related to not supporting all music consumption habits and browsing styles such as textual search or listening to full albums. These became even more important in the longer-term use.

Five novel music discovery prototypes using the same content-based music recommendation back-end were compared and the comparison results are reported, including first impressions and longer-term usage. The winner of this comparison having highest overall grade in longer-term use was the album cover space prototype (**P6**), offering innovative, simple and fast access to music. The album cover space enables the user to explore preview clips of different types of music by moving a finger around the screen. When a suitable preview clip has been found, the clip can seed an automatically generated playlist. The prototypes in the comparison only supported new music discovery and did not include the ability to, e.g., search for a specific song or artist.

All but one of the proposed music discovery concepts were considered to work well for music discovery. The use of avatar characters and mood pictures are the most encouraging concepts for music discovery. The use of mood pictures allows categorizing music beyond the traditional genre boundaries and including user-generated content, it provides for a rich music discovery experience. The avatars work well for visualizing playlists, music listening history, and are able to provide a playful and entertaining music discovery experience.

Especially, the evolving avatars based on music listening history (**P11**) were highly appreciated.

Based on the individual prototype findings, the use of collaborative mood pictures and adaptive avatars (**P10** and **P11**) offer a better overall user experience compared to the other concepts due to their wider support for different music discovery modes. They offer basic music searching capabilities in addition to playful music discovery and therefore do not include big hindrances regarding the support for different music consumption habits and potential longer-term use. Additionally, combining physical activities to music discovery was seen as a potential approach for further development (MyTerritory, **P9**). The virtual world prototype (**P2**) was not found suitable for music discovery because it made the task too time consuming and cumbersome. Having music recommendations as the sole activity in a virtual world was not the right approach for music discovery.

The results show that visually interesting and playful music discovery solutions can inspire users and work well, especially for exploratory music discovery. Even very simple and visual implementations can be considered fun, entertaining and suitable for quick music discovery but the experience will not last without the support for both active and passive music discovery. The results indicate that, when well designed, visual music discovery services can replace traditional music discovery services and support different types of music discovery practices.

This thesis contributes a generalized set of 16 design implications for music discovery service developers, which has not been introduced before in such magnitude for visual and playful music discovery services.

A wide range of future work ideas are presented in the included papers and they are further enhanced in the discussion chapter ranging from individual prototype enhancements towards future opportunities incorporating a wide range of sensor data and creative use of multiple modalities. One practical future step includes researching the use of automatic instrument recognition data for visual music discovery prototypes and exploring the potential use cases for new discovery opportunities. A second alternative is to further enhance the use of mood pictures by extending the set of preset pictures and offering new visual means for mood playlist categorization and moderation. A third interesting topic is to extend the adaptive avatar research with a broader questionnaire regarding the performance of avatars in representing both genres and eras, followed up with a common API for creating avatar representations for any playlists to be used across different services. Lastly, the concepts with the most potential should be evaluated with longer user trial periods of several months with larger music catalogs to learn more about the longer-term user experience and how well the concepts support larger amounts of content.

References

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. Bainbridge, W. *Encyclopedia of Human-Computer Interaction*. Thousand Oaks: Sage Publications, 37(4), 445-456.

Alonso, M. A., Richard, G. & David, B. (2004). Tempo And Beat Estimation Of Musical Signals. In ISMIR'04: *Proceedings of the 5th International Conference on Music Information Retrieval*. ISMIR.

Ankolekar, A., & Sandholm T. (2011). Foxtrot: a soundtrack for where you are. In *Proceedings of Interacting with Sound Workshop: Exploring Context-Aware, Local and Social Audio Applications* (pp. 26-31). ACM Press.

Arhippainen, L., & Hickey, S. (2011). Classifying music user groups and identifying needs for mobile virtual music services. In MINDTREK '11: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 191-196). ACM Press.

Arrasvuori, J., Korhonen, H., & Väänänen-Vainio-Mattila, K. (2010). Exploring playfulness in user experience of personal mobile products. In OzCHI '10: *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction* (pp. 88-95). ACM Press.

Arrasvuori, J., Boberg, M., Holopainen, J., Korhonen, H., Lucero, A., & Montola, M. (2011). Applying the PLEX framework in designing for playfulness. In DPPI '11: *Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces* (Article 24, 8 pages). ACM Press.

Averbukh, V., Bakhterev, M., Baydalin, A., Ismagilov, D., & Trushenkova, P. (2007). Interface and visualization metaphors. In HCII '07: *Proceedings of Human-Computer Interaction. Interaction Platforms and Techniques* (pp. 13-22). Springer Berlin Heidelberg.

Battarbee, K. (2004). *Co-experience, understanding user experiences in social interaction.* Doctoral dissertation. University of Art and Design, Helsinki.

Bentley, F., Metcalf, C., & Harboe, G. (2006). Personal vs. commercial content: the similarities between consumer use of photos and music. In CHI '06: *Proceedings of the 24th International Conference on Human Factors in Computing Systems* (pp. 667-676). ACM Press.

Bergman, J., Kauko, J., & Keränen, J. (2009). Hands on music: physical approach to interaction with digital music. In MobileCHI '09: *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM Press.

Boberg, M., Piippo, P., & Ollila, E. (2008). Designing avatars. In DIMEA '08: *Proceedings of 3rd International Conference on Digital Interactive Media in Entertainment and Arts* (pp. 232-239). ACM Press.

Bogdanov, D., Haro, M., Fuhrmann, F., Xambó, A., Gómez, E., & Herrera, P. (2013). Semantic audio content-based music recommendation and visualization based on user preference examples. *Information Processing & Management*, 49(1), 13-33.

Bossard, L., Kuhm, M., & Wattenhofer, R. (2009). Visually and Acoustically Exploring the High-Dimensional Space of Music. In CSE '09: *Proceedings on the 2009 International Conference on Computational Science and Engineering* (pp. 58-65). IEEE Computer Society.

Brahnam, S. (2001). Creating physical personalities for agents with faces: Modelling trait impressions of the face. In UM '01: *User Modelling 2001, workshop on Attitudes, Personality and Emotions in User-Adapted Interactions*, retrieved April 7, 2014 from http://aos2.uniba.it:8080/sonthofen/brahnam-f.pdf.

Brown, B., Geelhoed, E., & Sellen, A. (2001). The use of conventional and new music media: Implications for future technologies. In INTERACT ’01: *Proceedings of Interact 2001* (pp. 67-75). ACM press.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4), 331–370.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Casey, M. A., Veltkamp, R., Goto, M., Leman, M., Rhodes, C., & Slaney, M. (2008). Content-based music information retrieval: Current directions and future challenges. *Proceedings of the IEEE*, 96(4), 668-696.

Celma, O. (2008). *Music recommendation and discovery in the long tail*. Doctoral dissertation. Universitat Pompeu Fabra, Barcelona.

Celma, O., & Herrera, P. (2008). A new approach to evaluating novel recommendations. In RecSys ’08: *Proceedings of the 2008 ACM conference on Recommender systems* (pp. 179-186). ACM Press.

Corthaut, N., Govaerts, S., Verbert, K., & Duval, E. (2008). Connecting the dots: Music metadata generation, schemas and applications. In ISMIR’08: *Proceedings of the 9th International Conference on Music Information Retrieval* (pp. 249-254). ISMIR.

Creswell, J. W. (2008). *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, New Jearsey: Pearson.

Cunningham, S. J., Reeves, N., & Britland, M. (2003). An ethnographic study of music information seeking: implications for the design of a music digital library. In *Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital libraries* (pp. 5-16). IEEE Computer Society.

Cunningham, S.J., Bainbridge, D., & McKay, D. (2007). Finding new music: a diary study of everyday encounter with novel songs. In ISMIR ’07: *Proceedings of the 8th International Conference on Music Information Retrieval* (pp. 83-88), ISMIR.

Desmet, P. M. A. (2003). Measuring emotions. In M. A. Blythe, K. Overbeeke, A. F. Monk & P. C. Wright (Eds.). *Funology: From Usability to Enjoyment*. Dordrecht: Kluwer Academic Publishers.

Dey, A. K. (2001). Understanding and Using Context. *Personal and Ubiquitous Computing*, 5(1), 4-7, Springer-Verlag.

Dixon, S., & Cambouropoulos, E. (2000). Beat tracking with musical knowledge. In ECAI ’00: *Proceedings of the 8th European Conference on Artificial Intelligence* (pp. 626-630), Pitman.

Dixon, S. (2001). Automatic extraction of tempo and beat from expressive performances. *Journal of New Music Research*, 30(1), 39-58, Routledge.

Downie, J. S. (2003). Music information retrieval. Annual Review of Information Science and Technology, 37, 295-340. Available from http://music-ir.org/downie_mir_arist37.pdf.

Dunker, P., Nowak, S., Begau, A., & Lanz, C. (2008). Content-based mood classification for photos and music: a generic multi-modal classification framework and evaluation approach. In ICMR ’08: *Proceedings of the 1st ACM international conference on Multimedia information retrieval* (pp. 97-104). ACM Press.

Ekman, P. (2004). *Emotions revealed – Understanding faces and feelings*. London: Orion Books.

El Refaie, E. (2003). Understanding visual metaphor: The example of newspaper cartoons. *Visual Communication*, 2(1), 75-95.

Ellis, K., Coviello, E., Chan, A.B., & Lanckriet, G. (2013). A bag of systems representation for Music auto-tagging. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(12), 2554-2569, IEEE Signal Processing Society.

Eronen, A. (2001). Comparison of features for musical instrument recognition. In *IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics* (pp. 19-22), IEEE.

Eronen, A.J., Peltonen, V.T., Tuomi, J.T., Klapuri, A.P., Fagerlund, S., Sorsa, T., Lorho, G., & Huopaniemi, J. (2006). Audio-based context recognition. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(1), 321-329, IEEE Signal Processing Society.

Eronen, A. (2009). *Signal processing methods for audio classification and music content analysis*. Doctoral dissertation. Tampere University of Technology, Tampere.

Forceville, C. (1994). Pictorial metaphor in advertisements, *Metaphor and Symbolic Activity* 9(1), 1-29.

Fuhrmann, F. (2012). *Automatic musical instrument recognition from polyphonic music audio signals*. Doctoral dissertation, Universitat Pompeu Fabra, Barcelona.

Giannoulis, D., & Klapuri, A. (2013). Musical instrument recognition in polyphonic audio using missing feature approach. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(9), 1805-1817.

Glaser, B. (1992). *Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.

Goddard, C. (2011). *Semantic analysis: A practical introduction*. Oxford University Press.

Goto, M., & Goto, T. (2005). Musicream: New music playback interface for streaming, sticking, sorting, and recalling musical pieces. In ISMIR ’05: *Proceedings of the 6th International Conference on Music Information Retrieval* (pp. 404-411). ISMIR.

Haake, M., & Gulz, A. (2008). Visual stereotypes and virtual pedagogical agents. *Journal of Educational Technology & Society*, 11(4).

Hamel, P., & Eck, D. (2010). Learning Features from Music Audio with Deep Belief Networks. In ISMIR ’10: *Proceedings of the 11th International Conference on Music Information Retrieval* (pp. 339-344), ISMIR.

Haro, M., Xambó, A., Fuhrmann, F., Bogdanov, D., Gómez, E., & Herrera, P. (2010). The musical avatar: A visualization of musical preferences by means of audio content description. In AM ’10: *Proceedings of the 5th Audio Mostly Conference* (Article 14). ACM Press.

Hassenzahl, M., Platz, A., Burmester, M., & Lehner, K. (2000). Hedonic and ergonomic quality aspects determine a software's appeal. In CHI’00: *Proceedings of the 18th International Conference on Human Factors in Computing Systems* (pp. 201-208). ACM Press.

Hassenzahl, M. (2003). The think and I: Understanding the relationship between user and product. In M. A. Blythe, K. Overbeeke, A. F. Monk & P. C. Wright (Eds.). *Funology: From Usability to Enjoyment.* Dordrecht: Kluwer Academic Publishers.

Hassenzahl, M. (2008). User experience (UX): towards an experiential perspective on product quality. In IHM ’08: *Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine* (pp. 11-15). ACM Press.

Herlocker, J. L., Konstan, J. A., & Riedl, J. (2000). Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work* (pp. 241-250). ACM Press.

Hekkert, P. (2006). Design aesthetics: principles of pleasure in design. *Psychology science,* 48(2), 157.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75-105.

Hiipakka, J., & Lorho, G. (2003). A spatial audio user interface for generating music playlists. In ICAD ’03: *Proceedings of the 2003 International Conference on Auditory Display* (pp. 267-270).

Hoashi, K., Hamawaki, S., Ishizaki, H., Takishima, Y., & Katto, J. (2009). Usability evaluation of visualization interfaces for content-based music retrieval systems. In ISMIR ’09: *Proceedings of the 10th International Conference on Music Information Retrieval* (pp. 207-212). ISMIR.

Holm, J., Aaltonen, A., & Siirtola, H. (2009). Associating colours with musical genres. *Journal of New Music Research*, 38(1), 87-100. Routledge.

Holm, J., Siirtola, H., & Laaksonen, L. (2010). Associating avatars with musical genres. In IV '10: *Proceedings of the 14th International Conference on Information Visualisation* (pp. 186-193), IEEE Computer Society.
Holm, J. (2012). *Visualizing music collections based on metadata: Concepts, user studies and design implications*. Doctoral dissertation. Tampere University of Technology, Tampere.
Hu, X., Bay, M., & Downie, J. S. (2007). Creating a simplified music mood classification ground-truth set. In ISMIR '07: *Proceedings of the 8th International Conference on Music Information Retrieval* (pp. 309-310). ISMIR.
Hu, X., Downie, J. S., & Ehmann, A. F. (2009). Lyric text mining in music mood classification. In ISMIR '09: *Proceedings of the 10th International Conference on Music Information Retrieval* (pp. 411-416). ISMIR.
Hu, X., & Liu, J. (2010). Evaluation of music information retrieval: Towards a user-centered approach. In HCIR '10: *Proceedings of the 4th Workshop on Human-Computer Interaction and Information Retrieval*. ACM Press.
Huber, S, Scheld, M, & Knees, P. (2012). nepDroid: An intelligent mobile music player. In ICMR '12: *Proceedings of the 2nd ACM International Conference on Multimedia Retrieval*. ACM Press.
Isaac, S., & Michael, W. B. (1971). *Handbook in research and evaluation*. San Diego: EdITS.
ISO 9241-210 (2010). Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. *Switzerland International Standardization Organization*.
Jokela, T., Iivari, N., Matero, J., & Karukka, M. (2003). The standard of user-centered design and the standard definition of usability: analyzing ISO 13407 against ISO 9241-11. In *Proceedings of the Latin American conference on Human-computer interaction* (pp. 53-60). ACM Press.
Jones, M., & Marsden, G. (2006). *Mobile interaction design*. Chichester: John Wiley & Sons.
Julià, C. F., & Jorda, S. (2009). SongExplorer: A tabletop application for exploring large collections of songs. In ISMIR '09: *Proceedings of the 10th International Conference on Music Information Retrieval* (pp. 675-680). ISMIR.
Juslin, P. N., & Sloboda, J. A. (2001). *Music and emotion: Theory and research*. Oxford University Press.
Karapanos, E., Zimmerman, J. F., Forlizzi, J., & Martens, J-B. (2009). User experience over time: an initial framework. In CHI '09: *Proceedings of the 27th international conference on Human factors in computing systems* (pp. 729-738). ACM Press.
Karray, F., Alemzadeh, M., Saleh, J. A., & Arab, M. N. (2008). Human-computer interaction: Overview on state of the art. *International Journal on Smart Sensing and Intelligent Systems*, 1(1), 137-159.
Keränen, J., Bergman, J., & Kauko, J. (2009). Gravity sphere: gestural audio-tactile interface for mobile music exploration. In CHI '09: *Proceedings of the 27th International Conference on Human Factors in Computing Systems* (pp. 1531-1534). ACM Press.
Kim, Y., Schmidt, E., & Emelle, L. (2008). Moodswings: A collaborative game for music mood label collection. In ISMIR'08: *Proceedings of the 9th International Conference on Music Information Retrieval* (pp. 549–554). ISMIR.
Kim, H-J., Yoo, M-J., Kwon, J-Y., & Lee, I-K. (2009). Generating affective music icons in the emotion plane. In CHI'09: *Proceedings of the 27th International Conference on Human Factors in Computing Systems* (pp. 3389-3394). ACM Press.
Kim, Y. E., Schmidt, E. M., Migneco, R., Morton, B. G., Richardson, P., Scott, J., Speck, J., & Turnbull, D. (2010). Music emotion recognition: A state of the art review. In ISMIR '10: *Proceedings of the 11th International Conference on Music Information Retrieval* (pp. 255-266). ISMIR.
Knees, P., Schedl, M., Pohle, T., & Widmer, G. (2006). An innovative three-dimensional user interface for exploring music collections enriched. In ACM MM '06: *Proceedings*

of the 14th annual ACM international conference on Multimedia (pp. 17-24), ACM Press.
Konstan, J. A., & Riedl, J. (2012). Recommender systems: from algorithms to user experience. *User Modeling and User-Adapted Interaction*, 22(1-2), 101-123.
Korhonen, H., Montola, M., & Arrasvuori, J. (2009). Understanding playful user experience through digital games. In DPPI ’09: *Proceedings of the International Conference on Designing Pleasurable Products and Interfaces* (pp. 274-285).
Kujala, S., Vogel, M., Pohlmeyer, A. E., & Obrist, M. (2013). Lost in time: the meaning of temporal aspects in user experience. In CHI ’13: *Proceedings of the 31th International Conference on Human Factors in Computing Systems* (pp. 559-564). ACM Press.
Kuniavsky, M. (2003). *Observing the user experience: a practitioner's guide to user research*. Morgan Kaufmann.
Lacey, A., & Luff, D. (2001). *Qualitative data analysis*. Sheffield: Trent Focus.
Laplante, A., & Downie, J. S. (2006). Everyday life music information-seeking behaviour of young adults. In ISMIR ’06: *Proceedings of the 7th International Conference on Music Information Retrieval* (pp. 381-382), ISMIR.
Lehikoinen, J., Aaltonen, A., Huuskonen, P., & Salminen, I. (2007). *Personal content experience: managing digital life in the mobile age*. John Wiley & Sons.
Lehtiniemi, A., & Seppänen, J. (2007). Evaluation of automatic mobile playlist generator. In MOBILITY ’07: *Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology* (pp. 452-459). ACM Press.
Leppänen, J., & Eronen, A. (2013). Accelerometer-based activity recognition on a mobile phone using cepstral features and quantized gmms. In ICASSP ’13: *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 3487-3491), IEEE Signal Processing Society.
Lester, J., Choudhury, T., & Borriello, G. (2006). A practical approach to recognizing physical activities. In PERVASIVE ’06: *Proceedings of the 4th international conference on Pervasive Computing* (pp. 1-16), Springer-Verlag.
Lillie, A.S. (2008). *MusicBox: Navigating the space of your music.* BSc thesis, Stanford University.
Manning, C. D., Raghavan, P., & Schütze, H. (2008*). Introduction to information retrieval (Vol. 1)*. Cambridge: Cambridge university press.
Mao, J-Y., Vredenburg, K., Smith, P. W., & Carey, T. (2005). The state of the user-centered design practice. *Communications of the ACM, 48(3)*, 105-109.
McKay, C., & Fujinaga, I. (2006). Musical genre classification: Is it worth pursuing and how can it be improved?. In ISMIR ’06: *Proceedings of the 7th International Conference on Music Information Retrieval* (pp. 101-106). ISMIR.
Merčun, T., & Žumer, M. (2010). Visualizing for explorations and discovery. In LIDA ’10: *Proceedings of the international conference on Libraries in the Digital Age* (pp. 1-11).
Ming, L., Rozgić, V., Thatte, G., Sangwon L., Emken, B.A., Annavaram, M., Mitra, U., Spruijt-Metz, D., & Narayanan, S. (2010). Multimodal Physical Activity Recognition by Fusing Temporal and Cepstral Information. In *Neural Systems and Rehabilitation Engineering, IEEE Transactions*, 18(4), 369-380, IEEE Engineering in Medicine and Biology Society.
Mitrović, D., Zeppelzauer, M., & Breiteneder, C. (2010). Features for content-based audio retrieval. *Advances in computers*, 78, 71-150. Elsevier.
Nielsen, J. (1994). *Usability engineering*. Elsevier.
Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things.* New York, NY: Basic Books.
Okada, K., Karlsson, B. F., Sardinha, L., & Noleto, T. (2013). ContextPlayer: learning contextual music preferences for situational recommendations. In *Proceedings of the SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications.* ACM Press.
Orio, N. (2006). *Music retrieval: A tutorial and review (Vol. 1*). Now publishers Inc.

Pampalk, E. (2001). *Islands of Music: Analysis, Organization, and Visualization of Music Archives*. Master's thesis, Vienna University of Technology, 2001. Available from: http://www.oefai.at/~elias/music/thesis.html.

Patton, M.Q. (2002). *Qualitative Research and Evaluation Methods*. Thousand Oaks, CA: Sage.

Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. In *Artificial Intelligence Review*, 13(5-6), 393–408.

Peeters, G. (2004). A large set of audio features for sound description (similarity and classification) in the CUIDADO project. *CUIDADO Project Report*, Available from: http://recherche.ircam.fr/anasyn/peeters/ARTICLES/.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, 84(6), 1236–1256.

Russell, J. A (1980). A Circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161-1178.

Salber, D., Dey, A. K., & Abowd, G. D. (1999). The context toolkit: aiding the development of context-enabled applications. In CHI 99: *Proceedings of the 17th International Conference on Human Factors in Computing Systems* (pp. 434-441). ACM Press.

Schedl, M. (2008). *Automatically extracting, analyzing, and visualizing information on music artists from the World Wide Web*. Doctoral dissertation. Johannes Kepler University Linz, Austria.

Schedl, M., Flexer, A., & Urbano, J. (2013). The neglected user in music information retrieval research. *Journal of Intelligent Information Systems*, 41(3), 523-539.

Schneider, D. J. (2005). *The psychology of stereotyping*. Guilford Press.

Serra, X. (2011). A Multicultural Approach in Music Information Research. In ISMIR '11: *Proceedings of the 12th International Conference on Music Information Retrieval* (pp. 151-156).

Skowronek, J., McKinney, M. F., & Van De Par, S. (2006). Ground truth for automatic music mood classification. In ISMIR '06: *Proceedings of the 7th International Conference on Music Information Retrieval* (pp. 395-396), ISMIR.

Sloboda, J. A., & Juslin, P.N. (2001). Psychological perspectives on music and emotion. In P.N. Juslin & J. A. Sloboda (Eds.). *Music and emotion: Theory and research.* New York, NY: Oxford University Press.

Stober, S. (2011). *Adaptive methods for user-centered organization of music collections*. Doctoral dissertation. Otto-von-Guericke-University, Magdeburg, Germany.

Strauss, A. & Corbin, J. (1994). *Grounded theory methodology*. An Overview Handbook of Qualitative Research. Handbook of Qualitative Research, 273-285.

Sturm, B. L. (2012). Two systems for automatic music genre recognition: What are they really recognizing?. In MIRUM '12: *Proceedings of the second international ACM workshop on Music information retrieval with user-centered and multimodal strategies* (pp. 69-74). ACM Press.

Trohidis, K., Tsoumakas, G., Kalliris, G., & Vlahavas, I. (2008). Multi-label classification of music into emotions. In *ISMIR'08: Proceedings of the 9th International Conference on Music Information Retrieval* (pp. 325-330). ISMIR.

Thayer, R. E. (1989). *The biopsychology of mood and arousal*. Oxford University Press.

Turnbull, D., Barrington, L., Torres, D., & Lanckriet, G. (2008). Semantic annotation and retrieval of music and sound effects. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(2), 467-476. IEEE.

Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals., *IEEE Transactions on Speech and Audio Processing*, 10(5), 293-302. IEEE.

Tzanetakis, G., Kapur, A., Schloss, W. A., & Wright, M. (2007). Computational ethnomusicology. *Journal of interdisciplinary music studies*, 1(2), 1-24.

Van Aken, J. E. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and ground technological rules. *Journal of management Studies*, 41(2), 219-246.

Van Gulik, R., & Vignoli, F. (2005). Visual playlist generation on the artist map. In ISMIR ’05: *Proceedings of the 6th International Conference on Music Information Retrieval* (pp. 520-523). ISMIR.

Veletsianos, G. (2010). Contextually relevant pedagogical agents: Visual appearance, stereotypes, and first impressions and their impact on learning. *Computers & Education*, 55(2), 576-585.

Vignoli, F. (2004). Digital Music Interaction Concepts: A User Study. In ISMIR ‘04: *Proceedings of the 5th International Conference on Music Information Retrieval*, ISMIR.

Voong, M., & Beale, R. (2007). Music organisation using colour synaesthesia. In CHI '07: *Proceedings of the 25th international conference on Human factors in computing systems* (pp. 1869-1874). ACM Press.

Väänänen-Vainio-Mattila, K., Wäljas, M., Ojala, J., & Segerståhl, K. (2010). Identifying drivers and hindrances of social user experience in web services. In CHI ’10: *Proceedings of the 28th international conference on Human factors in computing systems* (pp. 2499-2502). ACM Press.

Wang, A. (2006). The Shazam music recognition service. In *Communications of the ACM, 49(8),* 44-48. ACM Press.

Wang, X., Rosenblum, D., & Wang, Y. (2012). Context-aware mobile music recommendation for daily activities. In MM ‘12: *Proceedings of the 20th ACM international conference on Multimedia* (pp. 99-108). ACM Press.

Weigl, D. M., & Guastavino, C. (2011). User Studies in the Music Information Retrieval Literature. In ISMIR ’11: *Proceedings of the 12th International Society for Music Information Retrieval Conference* (pp. 335-340). ISMIR.

Wiklund-Engblom, A., Hassenzahl, M., Bengs, A., & Sperring, S. (2009). What needs tell us about user experience. In INTERACT ’09: *Human-Computer Interaction* (pp. 666-669). Springer Berlin Heidelberg.

Yu, Y., & Nam, T. J. (2014). Let's giggle!: design principles for humorous products. In DIS 14: *Proceedings of the 2014 conference on Designing interactive systems* (pp. 275-284). ACM Press.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, 8(4), 494.

Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. *Applications of social research methods to questions in information and library science*, 308-319.

Web references

8tracks (2008). 8tracks web pages. Retrieved March 7, 2014 from http://www.8tracks.com

AllMusic (1991). All Music Guide. Retrieved March 31, 2014 from http://www.allmusic.com

Beats music (2014). Beats music web pages. Retrieved March 17, 2014 from http://www.beatsmusic.com

Bloom.fm (2013). Bloom.fm web pages. Retrieved March 7, 2014 from http://www.bloom.fm

Bylin, K. (2013). For Music Discovery, Definition Often Varies. Retrieved March 7, 2014 from http:// http://www.hypebot.com/hypebot/2013/04/music-discovery.html

Castranova, E. (2003). Theory of the avatar. Retrieved March 7, 2014 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=385103

Deezer (2007). Deezer web pages. Retrieved March 7, 2014 from http.//www.deezer.com

Discogs (2000). Wikipedia page on Discogs. Retrieved July 30, 2014 from http://en.wikipedia.org/wiki/Discogs

Discovr (2011). Wikipedia page on Discovr. Retrieved March 7, 2014 from http://en.wikipedia.org/wiki/Discovr

Forgotify (2013). Forgotify web pages. Retrieved March 11, 2014 from http://www.forgotify.com

Echonest (2005). The Echo Nest web pages. Retrieved March 14, 2014 from http://www.echonest.com

Freebase (2007). Freebase web pages. Retrieved March 6, 2014 from http://www.freebase.com

Google Streetview (2014). Google Streetview as part of Google Maps. Retrieved March 14, 2014, from https://www.google.com/maps/views/streetview

Gracenote (2000). Gracenote web pages. Retrieved March 13, 2014 from https://www.gracenote.com

Grooveshark (2007). Grooveshark web pages. Retrieved March 7, 2014 from http.//www.grooveshark.com

Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G., & Verplank. B. (1992). ACM SIGCHI Curricula for Human-Computer Interactin, Section 2.1: Definition of HCI. Retrieved March 24, 2014 from http://old.sigchi.org/cdg/cdg2.html#2_1

ID3 (1998). ID3 web pages, Retrieved March 13, 2014 from http://www.id3.org

IFPI (2014). Digital Music Report 2014, Retrieved March 25, 2014 from http://www.ifpi.org/downloads/Digital-Music-Report-2014.pdf

iTunes Genius (2008). Wikipedia information regarding iTunes Genius. Retrieved March 7, 2014 from http://en.wikipedia.org/wiki/ITunes_Genius

KCRW Music Mine (2013). KCRW Music Mine web pages. Retrieved March 7, 2014 from http://www.kcrw.com/about/musicmine-for-iphone

Last.fm (2002). Last.fm web pages. Retrieved March 6, 2014 from http://www.last.fm

Levitin, D. J. (2007). Life Soundtracks: The Uses of Music in Everyday Life. Retrieved March 7, 2014 from http://www.russballard.com/rbv7-workshop/physics/Unit%203%20Vibrations/LifeSoundtracks.pdf

Liveplasma (2005). Liveplasma web pages. Retrieved March 7, 2014 from http://www.liveplasma.com

MIREX (2013). Music Information Retrieval Evaluation eXchange 2013 results. Retrieved March 14, 2014 from http://www.music-ir.org/nema_out/mirex2013/results/act/mixed_report/

MIREX (2014). Music Information Retrieval Evaluation eXchange 2014 homepage. Retrieved July 30, 2014 from http://www.music-ir.org/mirex/wiki/MIREX_HOME

Moodagent (2001). Moodagent web pages. Retrieved March 6, 2014 from http://www.moodagent.com
Musicovery (2011). Musicovery web pages. Retrieved March 7, 2014 from http://musicovery.com
MusicBrainz (1999). MusicBrainz web pages. Retrieved March 6, 2014 from http://musicbrainz.org
Nokia Mixradio (2011). Nokia MixRadio web pages. Retrieved March 6, 2014 from http://www.nokia.com/mixradio/
Oxford (2014). Metaphor in Oxford dictionary. Retrieved July 31, 2014 from http://www.oxforddictionaries.com/definition/english/metaphor
Pandora (2000). Pandora radio web pages. Retrieved March 6, 2014 from http://www.pandora.com
Rdio (2010). Rdio web pages. Retrieved March 7, 2014 from http://www.rdio.com
Seevl (2012). Seevl web pages. Retrieved March 6, 2014 from http://seevl.fm
Schoenwald, J., & Lead, U. D. (2012). White paper solving the discovery problem. Retrieved April 9, 2014 from http://info.mportal.com/hs-fs/hub/134079/file-240327195-pdf/Whitepaper_Folder/mPortal_Discovery_2012.pdf
Shazam (1999). Shazam web pages. Retrieved March 7, 2014 from http://www.shazam.com
Shuffler.fm (2010). Shuffler.fm web pages. Retrieved March 7, 2014 from http://www.shuffler.fm
Sonarflow (2012). Sonarflow web pages. Retrieved March 7, 2014 from http://www.sonarflow.com/
Songza (2007). Songza web pages. Retrieved March 7, 2014 from http://www.songza.com
Songza (2014). Songza taps Weather Channel data to suggest mood-enhancing music. Retrieved April 1, 2014 from http://www.engadget.com/2014/03/27/songza-weather-channel/
Sony SensMe (2014). Sony SensMe Wikipedia page. Retrieved March 20, 2014 from http://en.wikipedia.org/wiki/SensMe
SoundCloud (2007). SoundCould web pages. Retrieved March 7, 2014 from http://www.soundcloud.com
SoundHound (2009). SoundHound web pages. Retrieved March 7, 2014 from http://www.soundhound.com
Soundwave (2011). Soundwave web pages. Retrieved March 7, 2014 from http://www.soundwave.com
Spotify (2006). Spotify web pages. Retrieved March 6, 2014 from http://www.spotify.com
Spotify introduction (2011). Spotify introduction. Retrieved April 2, 2014 from http://www.slideshare.net/ewilko/remarkable-spotify
Stereogum (2002). Stereogum web pages. Retrieved March 7, 2014 from http://www.stereogum.com
StereoMood (2012). Stereomood web pages. Retrieved March 7, 2014 from http://www.stereomood.com
TasteKid (2009). TasteKid web pages. Retrieved March 7, 2014 from http://tastekid.com
The Hype Machine (2005). The Hype Machine web pages. Retrieved March 6, 2014 from http://hypem.com
This is my jam (2013). This is my jam web pages. Retrieved March 7, 2014 from http://www.thisismyjam.com
Youtube (2005). Youtube web pages. Retrieved March 6, 2014 from http://www.youtube.com
Wikipedia (2001). Wikipedia web pages. Retrieved March 6, 2014 from http://www.wikipedia.org

Original publications

Tampereen teknillinen yliopisto
PL 527
33101 Tampere

Tampere University of Technology
P.O.B. 527
FI-33101 Tampere, Finland

ISBN 978-952-15-3392-1
ISSN 1459-2045