

Interactive Sonification of Motion

Design, implementation and control of expressive auditory feedback with mobile devices

GAËL DUBUS



Doctoral Thesis in
Sound and Music Computing
Stockholm, Sweden 2013



**KTH Computer Science
and Communication**



KTH Computer Science
and Communication

Interactive sonification of motion

Design, implementation and control of expressive
auditory feedback with mobile devices

GAËL DUBUS

Doctoral Thesis
Stockholm, Sweden 2013

TRITA-CSC-A 2013:09
ISSN 1653-5723
ISRN KTH/CSC/A-12/09-SE
ISBN 978-91-7501-858-4

KTH School of Computer Science
and Communication
SE-100 44 Stockholm
SWEDEN

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen i datalogi fredagen den 27 september 2013 klockan 13.00 i F3, Sing Sing, Kungl Tekniska högskolan, Lindstedtsvägen 26, Stockholm.

© Gaël Dubus, September 2013

Tryck: Universitetsservice US AB

Abstract

Sound and motion are intrinsically related, by their physical nature and through the link between auditory perception and motor control. If sound provides information about the characteristics of a movement, a movement can also be influenced or triggered by a sound pattern. This thesis investigates how this link can be reinforced by means of interactive sonification. Sonification, the use of sound to communicate, perceptualize and interpret data, can be used in many different contexts. It is particularly well suited for time-related tasks such as monitoring and synchronization, and is therefore an ideal candidate to support the design of applications related to physical training. Our objectives are to develop and investigate computational models for the sonification of motion data with a particular focus on expressive movement and gesture, and for the sonification of elite athletes movements. We chose to develop our applications on a mobile platform in order to make use of advanced interaction modes using an easily accessible technology. In addition, networking capabilities of modern smartphones potentially allow for adding a social dimension to our sonification applications by extending them to several collaborating users. The sport of rowing was chosen to illustrate the assistance that an interactive sonification system can provide to elite athletes. Bringing into play complex interactions between various kinematic and kinetic quantities, studies on rowing kinematics provide guidelines to optimize rowing efficiency, e.g. by minimizing velocity fluctuations around average velocity. However, rowers can only rely on sparse cues to get information relative to boat velocity, such as the sound made by the water splashing on the hull. We believe that an interactive augmented feedback communicating the dynamic evolution of some kinematic quantities could represent a promising way of enhancing the training of elite rowers. Since only limited space is available on a rowing boat, the use of mobile phones appears appropriate for handling streams of incoming data from various sensors and generating an auditory feedback simultaneously.

The development of sonification models for rowing and their design evaluation in offline conditions are presented in Paper I. In Paper II, three different models for sonifying the synchronization of the movements of two users holding a mobile phone are explored. Sonification of expressive gestures by means of expressive music performance is tackled in Paper III. In Paper IV, we introduce a database of mobile applications related to sound and music computing. An overview of the field of sonification is presented in Paper V, along with a systematic review of mapping strategies for sonifying physical quantities. Physical and auditory dimensions were both classified into generic conceptual dimensions, and proportion of use was analyzed in order to identify the most popular mappings. Finally, Paper VI summarizes experiments conducted with the Swedish national rowing team in order to assess sonification models in an interactive context.

Acknowledgements

First and foremost, I am infinitely grateful to my main supervisor Roberto Bresin, who has always been understanding and supportive, and a great source of inspiration for my work. Your accommodating and easy-going attitude has always been appreciated, not least when I was about to drown in the huge amount of papers collected for the *bomba*.

I would particularly like to thank my co-supervisor Anders Friberg for his constant good mood and great sense of humor. You are an inspiration to all of us, especially when enchanting us with great musical performances.

Many thanks to Anders Askenfelt and Sten Ternström who manage the Sound and Music computing group in a professional yet positive and friendly way.

I would like to thank my present and past colleagues in the Sound and Music Computing group: my roommates Kjetil and Andreas, Marco, Gláucia, Laura, Gerhard, Johan, Svante, Kahl, Roxana, Anick, Erwin, Jan, Carina, Dong Li, Peter, Per-Magnus, Ludvig, Anders, and Guillaume.

Thanks to my friends and colleagues from the Speech group and from the Language unit for many interesting and fun conversations.

It has been a great pleasure to play music during these years. Special thanks my jazzband, the fantastic PromenadorQuestern that made me travel from Trondheim to Marrakech and many other places in between. Thanks to Gunnar Julin and his wonderful KTHAK, to Engelbrekts Vokalensemble, to Det Norske Korps, and not to forget that managing the fabulous Formantorkestern has been the most pleasant part of my department duties!

Merci à mes précieux et surpuissants HX2, votre amitié sans faille a toujours été le plus grand des réconforts.

Merci à François et Loïc sans qui la vie à Stockholm n'aurait pas été aussi emplie de joie et d'émotions.

Merci à Mamie, Maman, Papa, Johan, Ariane, et toute ma famille qui m'a continuellement encouragé, que ce soit depuis la France ou en venant fréquemment me rendre visite dans le Grand Nord.

Tack till *ma chérie* Laura för ditt stöd, dina kloka ord, din snällhet och din kärlek. Tack för många fantastiska resor genom världen och livet. Je t'aime!

Gaël Dubus
Stockholm, August 30th 2013

Contents

Acknowledgements	v
Contents	vi
Papers included in the thesis	ix
Related publications	xiii
1 Introduction	1
I From design to evaluation	5
2 Sound in interaction	7
2.1 Musical interaction	7
2.2 Sonification	8
3 Implementing interactive sonification on mobile devices	13
3.1 Going mobile	13
3.2 Tools for sound design and sound interaction	14
4 Results	17
4.1 Database of Android applications related to sound and music computing	17
4.2 Applications developed	17
4.3 Mapping data to sound	21
4.4 Evaluation results	23
5 Discussion	25
5.1 Contributions	25
5.2 Conclusions	28
Bibliography	31

<i>CONTENTS</i>	vii
-----------------	-----

II Included papers	35
Paper I	37
Paper II	71
Paper III	107
Paper IV	129
Paper V	145
Paper VI	225

Papers included in the thesis

The papers are listed in chronological order.

Paper I

Dubus, G. (2012). “Evaluation of four models for the sonification of elite rowing”. *Journal on Multimodal User Interfaces*, 5(3–4), 143–156.

Paper II

Varni, G., Dubus, G., Oksanen, S., Volpe, G., Fabiani, M., Bresin, R., Kleimola, J., Välimäki, V., & Camurri A. (2012). “Interactive sonification of synchronisation of motoric behaviour in social active listening of music with mobile devices.” *Journal on Multimodal User Interfaces*, 5(3–4), 157–173.

Giovanna Varni wrote the main part of the article, contributed to the application design and development, to the evaluation experiment design, to data collection, and performed the statistical analysis. Gaël Dubus contributed to the application design and development, to the evaluation experiment design, to data collection, and manuscript authoring (Sections 3 and 5.3). Sami Oksanen contributed to the application design and development, to data collection, and manuscript authoring (Section 5.1). Gualtiero Volpe contributed to the application design and development, to the evaluation experiment design, and to data collection. Marco Fabiani contributed to the application design and development, to the evaluation experiment design, and to data collection. Roberto Bresin contributed to the application design and development, and to the evaluation experiment design. Jari Kleimola contributed to the application design and development, and to data collection. Vesa Välimäki contributed to the application design and development, and to data collection. Antonio Camurri contributed to the application design and development.

Paper III

Fabiani, M., Bresin, R., & Dubus, G. (2012). “Interactive sonification of expressive hand gestures on a handheld device” *Journal on Multimodal User Interfaces*, 6(1), 49–57.

Marco Fabiani wrote the main part of the article, contributed to the application design and development, to the evaluation experiment design and data collection, and to the statistical data analysis. Roberto Bresin contributed to the statistical data analysis and manuscript authoring. Gaël Dubus contributed to the application design and development, and to the evaluation experiment design and data collection.

Paper IV

Dubus, G., Hansen, K. F., & Bresin, R. (2012). “An overview of sound and music applications for Android available on the market.” In *Proceedings of SMC 2012, 9th Sound and Music Computing Conference* (pp. 541–546). Copenhagen, Denmark.

Gaël Dubus contributed to the design of the study, created the application database, defined the classification categories, performed the statistical analysis, and wrote the main part of the article. Kjetil Falkenberg Hansen contributed to the design of the study, and manuscript authoring (Section 2). Roberto Bresin contributed to the design of the study.

Paper V

Dubus, G., & Bresin, R. (submitted) A systematic review of mapping strategies for the sonification of physical quantities. Submitted to *PLoS ONE*.

Gaël Dubus contributed to the design of the study, created the publication database, defined the classification spaces, conducted the analysis of the included publications to form the mapping database, contributed to the design of the statistical analysis of the mapping database, performed the statistical analysis, and wrote the main part of the article. Roberto Bresin contributed to the design of the study, to the design of the statistical analysis, and to manuscript authoring (Section 2.1).

Paper VI

Dubus, G., & Bresin, R. (submitted). Evaluation of a system for the sonification of elite rowing in an interactive context. Submitted to *Sports Engineering*.

Gaël Dubus designed and performed the experiments, conducted the analysis of the data, and wrote the article. Roberto Bresin contributed to the design of the experiments.

Related publications

Camurri, A., Volpe, G., Vinet, H., Bresin, R., Fabiani, M., Dubus, G., Maestre, E., Llop, J., Kleimola, J., Oksanen, S., Välimäki, V., & Seppanen, J. (2010). User-centric context-aware mobile applications for embodied music listening. In Daras, P., Mayora Ibarra, O. (Eds.), *User Centric Media* (pp. 21–30). Springer Berlin Heidelberg.

Dubus, G., & Bresin, R. (2010). Sonification of sculler movements, development of preliminary methods. In Bresin, R., Hermann, T., & Hunt, A. (Eds.), *Proceedings of ISon 2010, 3rd Interactive Sonification workshop* (pp. 39–43). Stockholm, Sweden.

Fabiani, M., Dubus, G., & Bresin, R. (2010). Interactive sonification of emotionally expressive gestures by means of music performance. In Bresin, R., Hermann, T., & Hunt, A. (Eds.), *Proceedings of ISon 2010, 3rd Interactive Sonification workshop* (pp. 113–116). Stockholm, Sweden.

Dubus, G., & Bresin, R. (2011). Sonification of physical quantities throughout history: a meta-study of previous mapping strategies. In Wersényi, G., & Worrall, D. (Eds.) *Proceedings of ICAD 2011, 17th International Conference on Auditory Display* (CD-ROM). Budapest, Hungary.

Fabiani, M., Dubus, G., & Bresin, R. (2011). MoodifierLive: interactive and collaborative music performance on mobile devices. In JENSENIUS, A. R., Tveit, A., Godøy, R. I., & Overholt, D. (Eds.), *Proceedings of NIME'11, 11th International Conference on New Interfaces for Musical Expression* (pp. 116–119). Oslo, Norway.

Hansen, K. F., Dubus, G., & Bresin, R. (2012). Using modern smartphones to create interactive listening experiences for hearing impaired. In Bresin, R., & Hansen, K. F. (Eds.) *Proceedings of Sound and Music Computing Sweden 2012*, ser. TMH-QPSR, 52(1), 42. KTH Royal Institute of Technology.

À mes grand-parents

Chapter 1

Introduction

Sound is intrinsically related to movement: the physical nature of acoustic waves, being the mechanical vibration of a medium, implies that sound production requires a movement to generate an excitatory pattern. In the introduction to the course *Sound in interaction* given at KTH, challenging the students to make the loudest possible sound without moving a single muscle always results in an awkward silence. Many dimensions of a sound can carry information characterizing its source, and humans and animals have learnt to associate meaning to some of these dimensions at a low cognitive level. For instance, a loud sound will probably be understood as potentially more dangerous than a sound that is barely audible, because it involves a vibration of larger amplitude probably caused by stronger forces acting at the location of the sound source. Similarly, the pitch of a sound can often be associated to the size of the corresponding sounding object, due to the fact that resonant frequencies of smaller objects are higher (Giordano and McAdams, 2006). Interestingly, the same association was found when investigating the optimal frequency of animal communication (Fletcher, 2004), meaning that the perception of the cries of animals can provide information about their size. These direct associations, based on characteristics of the sound source and on the physical laws governing its environment, are universal and deeply engraved in us from birth. Other types of associations are integrated at an early age: for auditory perception, as for every other sensory modality, the development of cognitive processes in infants is conditioned by the interaction with the environment (von Hofsten, 2009). By performing and repeating actions, we can learn how the sound occurring when hitting a drum differs from the one of breaking a glass, and to distinguish between a metallic object and a wooden object via their characteristic timbre.

Sonification, the use of sound to communicate, perceptualize and interpret data, makes use of the fact that we use hearing to extract information from distant phenomena or to examine properties of surrounding objects. It could be described as a subdomain of augmented reality using the auditory modality, revealing information that would be hidden otherwise through the use of well-defined associations

between data dimensions and auditory dimensions of the sonification display. These associations, called *mappings*, are either metaphorical — and have then to be learnt from scratch, e.g. by interacting with the sonification system — or consist in exaggerated implementations of natural associations, for example in the process of sound cartoonification (Rocchesso *et al.*, 2003; Rath, 2003).

Auditory perception does not only provide information about a distant motion, it is also closely linked to motor control. It has been shown that a particular group of cortical neurons located in Broca's area of the human brain, called mirror neurons, play a major role in coding representation of actions (Kosslyn *et al.*, 2001; Kohler *et al.*, 2002; Buccino *et al.*, 2006). Particularly sensitive to audiovisual stimuli, they show a similar activity when the subject performs an action, observes someone performing this action or simply perceives auditory or visual stimuli associated to this action. These neurons play a major role for motor functions related to speech production, but also to imitation learning in a broader perspective. In addition, it is known that humans are better to synchronize their movements with a rhythm provided in the auditory modality than in the visual modality (Repp and Penel, 2004).

In a clever experiment, Caramiaux *et al.* (2011) have shown that if participants recognized the sound source (e.g. everyday sounds), then they represented the perceived sound with gestures mimicking the action that had generated it. Otherwise (e.g. for artificial sounds) they used gestures representing time-varying auditory cues of the perceived sound. Numerous applications of more or less sophisticated couplings between sound and movement can be found throughout history. On Roman galleys, the *hortator* was in charge of giving the pace to the rowers and keeping them synchronized by striking a drum. In a more elaborate interaction, learning how to play a musical instrument implies to be placed in a sensorimotor loop, adapting the timing and magnitude of specific gestures to obtain the desired auditory feedback provided by the instrument. Another example of strong link between sound and movement is found in the intimate relationship between music and body motion in the context of dancing. A dancer interprets and complements the music by performing expressive gestures synchronously, often with the perspective of creating a type of social interaction.

The main objectives of this thesis work are to develop and investigate computational models for the sonification of motion data with a particular focus on expressive movement and gesture, and for the sonification of elite athletes movements. We chose to develop our applications on a mobile platform in order to make use of advanced interaction modes using an easily accessible technology. In addition, networking capabilities of modern smartphones potentially allow for adding a social dimension to our sonification applications by extending them to several collaborating users. The considered expressive gestures correspond therefore to situations where users hold a mobile phone in their hand. The sport of rowing was chosen to illustrate the assistance that an interactive sonification system can provide to elite athletes. Bringing into play complex interactions between various kinematic and kinetic quantities, studies on rowing kinematics provide guidelines to

optimize rowing efficiency, e.g. by minimizing velocity fluctuations around average velocity. However, rowers can only rely on sparse cues to get information relative to boat velocity, such as the sound made by the water splashing on the hull. We believe that an interactive augmented feedback communicating the dynamic evolution of some kinematic quantities could represent a promising way of enhancing the training of elite rowers. Since only limited space is available on a rowing boat, the use of mobile phones appears appropriate for handling streams of incoming data from various sensors and generating an auditory feedback simultaneously.

The development of sonification models for rowing and their design evaluation in offline conditions are presented in Paper I. In Paper II, three different models for sonifying the synchronization of the movements of two users holding a mobile phone are explored. Sonification of expressive gestures by means of expressive music performance is tackled in Paper III. In Paper IV, we introduce a database of mobile applications related to sound and music computing. An overview of the field of sonification is presented in Paper V, along with a systematic review of mapping strategies for sonifying physical quantities. Finally, Paper VI summarizes experiments conducted with the Swedish national rowing team in order to assess sonification models in an interactive context.

Part I

From design to evaluation

Chapter 2

Sound in interaction

2.1 Musical interaction

Expressive computer-controlled music performance

Research on expressive computer-controlled music performance has been conducted for many years, the main objective of this field being to improve the ways a computer can play music from a symbolic score. Playing a MIDI file on a computer is probably largely recognized as a painful experience for the majority of non-expert people. There are two reasons for that: the (usually) poor quality of the default synthesizer, and the way the computer renders the score — the *performance*. When providing a music sheet to a human performer, he will try to follow the notation the closest as possible, yet no musician can perform a piece of music in a perfectly regular way. In fact, the deviation from nominal values is what makes music performance interesting: we call a musician a virtuoso based on his technical skills, but also on his abilities to perform music inventing his own particular deviations for a number of performance parameters (e.g. tempo, phrasing). On the other hand, a computer can play a perfectly timed and balanced piece, which inevitably sounds extremely boring and non-human.

A research objective in the field of Sound and Music Computing is therefore to find ways to make a computer-controlled music performance sound “more human”. To this end, researchers at the Department of Speech, Music and Hearing of KTH have developed a system for introducing deviations in a meaningful way by combining different performance parameters. The KTH rule system for expressive music performance, described in Friberg *et al.* (2006), includes a large set of elementary rules governing combined variations of diverse performance parameters. An extensive set of rules has been implemented in the program *Director Musices*. Combining these elementary rules at a higher level enables the creation of more advanced effects, e.g. putting more energy in a performance. This has been accomplished and implemented in another program, called *pDM*, capable of applying and changing effects in real-time while playing a music file (Friberg, 2006).

A map of various emotional states have been set up in a two-dimensional space called the activity-valence space. The dimensions of this space represent the amount of energy associated to a given emotional state (e.g. low for sadness, high for happiness or anger), and the arousal associated to it (e.g. negative for sadness and anger, positive for happiness). Using special combinations of rules to cover the activity-valence space, the KTH rule system for expressive music performance provides the possibility to give an emotional character to a computer-controlled music performance.

Active music listening

Systems such as the set of rules of *Director Musices* can exploit the strong link between body motion and music performance when implemented in an interactive environment. While relating body gestures to a meaningful auditory feedback in the form of an expressive music performance, we can create a sensorimotor loop encouraging an active listening behavior. Active music listening based on expressive computer-controlled music performance has been extensively discussed in Marco Fabiani's doctoral thesis (Fabiani, 2011), where two applications designed for interactive manipulation of music performance have been presented: *PerMORFer* and *Moodifier*. While the former aimed at producing a high-quality audio rendering of computer-manipulated music performance, the latter focused on new ways of interacting with the performance in an active listening loop, using mobile devices as control platform. In the framework of the present thesis, we introduce *Moodifier* as a tool for performing the sonification of expressive body gestures.

2.2 Sonification

In this section we present a short overview of sonification, with a brief presentation of the concept and most usual techniques, as well as common domains of application. A more detailed overview of the field can be found in the first section of Paper V: we discuss the nature of sonification by starting from the successive definitions, then taking a historical perspective to introduce some famous sonification applications.

Definition

The definition of sonification is still disputed nowadays in the scientific community. Theory of sonification is a recurrent topic at the International Conference on Auditory Display (ICAD), where two apparently antagonist tendencies coexist: the need to define more clearly the boundaries of the field, therefore restricting the general character of the definition (Hermann, 2008), and the urge to expand these boundaries to include more artistic types of work and get closer to the world of music. Probably the most commonly accepted definition states that “*sonification is the use of non-speech audio to convey information*” (Walker and Nees, 2011). In the conclusion of Paper V, we observe that this formulation is confusing: it is

unclear whether the exclusion of speech should apply to the sonic material constituting the auditory display or to the auditory dimensions forming the codomain of the possible mappings. Considering that using speech sounds as sonic material should be allowed, for example using a mapping from a given input dimension to the loudness or pitch of a vowel, we point out the need for a reformulation of the definition. Based on observations originating from the classification process used in our systematic review, we suggest putting back the concept of mapping in the definition of sonification, as introduced in one of the earlier definitions by Scaletti (1994). In our study, we developed a characterization of sonification of physical quantities by examining if a mapping could be identified between a physical dimension and an auditory dimension. By extending this characterization to all types of input data, we can characterize sonification itself. Defining the boundaries of the field would be achieved by restricting the domain (the set of input data dimensions) and the codomain (the set of auditory dimensions) of the possible mappings. In the previous example, the exclusion of “speech” would then be understood as a limitation of the codomain of the possible mappings, excluding the *semantics* associated to words.

The most common techniques are audification, auditory icons, earcons, parameter mapping sonification, and model-based sonification. They are well documented in the *Sonification handbook* (Hermann *et al.*, 2011), the most recent publication providing an exhaustive overview of the field. They are also explained in the first Section of Paper V.

Applications

Although displaying data is, in general, performed more frequently via scientific visualization, some particular domains are particularly well suited to the use of sonification. Defined at first as its auditory counterpart, sonification can be used to complement visualization or to replace it in contexts where vision is unavailable, obstructed, or needs to be free from distractions (e.g. for a surgeon in an operating room). Due to the temporal nature of sound, time-related tasks can be easily enhanced by the use of an auditory feedback. As mentioned in the introduction to this thesis, movement, and especially human body motion, is naturally coupled to sound. Therefore, tasks such as synchronization of gestures (an example of which is presented in Paper II), sports training (Papers I and VI) and rehabilitation (Godbout and Boyd, 2010) could potentially benefit from a sonification system. A broad classification of the primary function of sonification systems in each research project was performed in Paper V (Fig. 4). The following categories emerged, ranked from most to least frequent: data exploration, art and aesthetics, accessibility, motion perception, monitoring, complement to visualization, and study of psychoacoustics. It could be observed that a significant proportion of projects were focused on art and aesthetics, showing that sonification is not limited to purely scientific applications. On the other hand, few projects had psychoacoustics as primary concern,

which illustrates the fact that evaluation of sonification systems is usually not the main matter of research yet.

Using music material to create sonification

In Section 2.1 we presented advanced combinations of different high-level parameters of music performance (e.g. articulation, tempo, phrasing). These associations of parameters result from studies based on auditory and musical perception and were designed to create expressive effects. Moreover, for aesthetic purposes, it is often tempting to use musical material in an auditory display, musical sounds being judged less intrusive are more easily accepted than other synthetic sounds commonly used in sonification (e.g. pure tones, noise bursts). For these two reasons, it appears appropriate to incorporate these associations in the design of sonification mappings.

However, the relationship between music and sonification is a controversial subject among sonification researchers. The reason for that seems to lie in the common perception of sonification by the public: when understanding that sonification is about creating organized sound patterns and is directed towards human perception, a layperson would often call it “music”. This is obviously erroneous as sonification is first and foremost a scientific discipline. In order to avoid such misunderstandings, many attempts were made to draw clear boundaries between sonification and music. A permanent feature in the successive definitions that were elaborated in this perspective is that the primary purpose of sonification is to communicate objective data, therefore excluding highly interpretive variables such as emotions from the set of possible inputs. At ICAD 2011, Emery Schubert, a musicologist, held a deliberately provocative talk, beginning as follows: *“I am here to teach you, ICAD people, what sonification is about. My claim is that music has the potential to sonify our emotions”*. If this claim goes fundamentally against the vision of the great majority of researchers and sonification pioneers, it illustrates the ambivalent relationship between the two domains: there is often an artistic dimension in sonification design, and music can sometimes be composed on the basis of scientific data, or in the perspective of communicating objective information.

In Section 3 of Paper II, we discuss the issue of using musical material in sonification, and we elaborate on the ways to discriminate between sonification and data-driven music. We conclude that the intention of the work is fundamental, the exact same sound track being potentially considered as sonification (if the purpose was to communicate information about a certain data set) — or music (if the project was rather motivated by aesthetic considerations). In Paper V we extend this view by making the distinction between the sonic material (that could possibly be based on music) and the mappings characterizing sonification.

Sonification applications presented in Papers II (*Sync’n’Mood*) and III (*MoodifierLive*) are both based on the KTH rule system for expressive music performance. In both cases, the most restrictive definitions such as the one proposed by Hermann (2008) would not classify these applications as sonification systems. Taking

a mapping-centered approach to define sonification, we can define high-level auditory dimensions such as the activity and valence of the performance, which are in fact combinations of four parameters: tempo, sound level, articulation, and phrasing. In the case of *Sync'n'Mood*, the input data dimension is a physical parameter based on acceleration data and reflecting the synchronization between two users. It is therefore easy to define a mapping characterizing the process as sonification of a physical quantity (the project was indeed considered in the systematic review conducted in Paper V). On the other hand, *ModifierLive* makes use of more indirect mappings, requiring to be particularly careful when qualifying it as sonification: it performs the sonification of expressive gestures based on acceleration data, not of emotions (even if the gestures are classified according to emotion labels). While apparently convoluted, this example of sonification was designed to enable advanced modes of interaction and trigger active music listening for the user.

Role of interaction

Interaction plays a major role in sonification applications, first of all at the design level where user behavior in relation to the system should be considered central. Applications that are defined as user-centric, i.e., when the experience of the users depends greatly on their behavior towards the system rather than on the specific features of the content (e.g. score, music genre, sound synthesis algorithm), encourage them to be active in their interaction with an auditory (or multimodal) feedback. The applications presented in Papers II and III, using music material as basis for the auditory display, were designed specifically to place users in an active listening mode. In addition, the three sonification models presented in Paper II added a social dimension to the experience by establishing a collaboration between the two participants, which had the effect of increasing their engagement. Following this principle, we proposed to extend the application presented in Paper III in order to create a local collaborative network, each participant being able to control a dimension of the expressive music performance (Fabiani *et al.*, 2011).

The level of interaction is also decisive for the acceptance of the system. In special contexts such as elite sports training, athletes are required to stay perfectly focused while in a demanding situation, therefore an intrusive sound would not be accepted. The design of a continuous auditory feedback is a challenging task, and an evaluation of prototypes of sonification via listening tests, i.e. in a non-interactive situation, can be used to get a better hindsight on aesthetic preferences of the potential users. Nevertheless, a higher acceptance rate can be expected after interactive experiments: as it was explained by Hunt *et al.* (2004), sounds that would be considered annoying otherwise turn out to be accepted by users of set in an interactive control loop with the auditory feedback system. Higher acceptance rates were indeed measured in Paper VI, for which interactive experiments were conducted, than in Paper I although the models were the same.

Challenges

As a relatively young field of research (it started to be studied as such in the beginning of the 1990s), sonification still needs to overcome many difficulties to become an established discipline. In Paper V, we try to identify some of these challenges, and to address them by conducting a systematic review of sonification works. One of the main difficulties is to organize the knowledge to be able to compare different studies in terms of design, goals, and technology. There is an obvious lack of unity in the terminology of the data spaces, both for input data dimensions and auditory display. To address this issue, we developed a flexible classification of dimensions used in sonification of physical quantities: the dimensions described in the considered research project, showing an extreme diversity, were gathered into conceptual intermediate-level dimensions such as velocity, temperature, density, or length. In this way, similar mappings could be compared, and we could determine the most popular associations. However, this classification was based on our understanding of the low-level dimensions, and incorporates therefore a certain level of subjectivity. Cooperation with other researchers from the International Community for Auditory Display is needed in order to reach a more objective and more stable classification, but there is obviously much to learn from other disciplines when it comes to input data dimensions. Another objective of Paper V was to list the mappings having been assessed via a rigorous evaluation process, either as successful or unsuccessful. The systematic review showed that, if a functional evaluation of sonification applications is often performed, the design of sonification mappings is still arbitrary in most of the cases. There is therefore a need for established evaluation methods based on psychoacoustical experiments to be performed more systematically. We suggest setting the focus on mappings that have often been used without having been assessed, thus requiring a particular attention — and, ideally, a psychoacoustical evaluation as performed by Walker (2007) — in future applications.

Chapter 3

Implementing interactive sonification on mobile devices

3.1 Going mobile

In this section we underline the specificities of a mobile platform, as compared to the usual programming environment of a stationary computer. We describe the numerous challenges encountered when working in this environment, as well as the novel opportunities brought by rapid advances in technology.

The technology race

Mobile technology appears to evolve at an ever-increasing pace: in the scope of fifteen years, dozens of generations of mobile devices have succeeded each other. At first designed exclusively to make phone calls, mobile phones have evolved to include more functions (e.g. pager, text messages, calendar, games) until the first generation of smartphones came, containing all the functions of a small computer. Both hardware and computational capabilities have improved greatly to support more and more advanced tasks. In addition, they could benefit from a diversification of accessible sensors (e.g. accelerometer, proximity sensor, GPS, gyroscope, microphone, touch screen, camera) to become computing artefacts aware of their context of use, able to connect to local and global social environments thanks to the development of networking capabilities (e.g. Bluetooth, wireless Internet, 3G).

A vision of what technology could bring in our everyday life, and what it already brings to us, is given by Rocchesso and Bresin (2007). The authors focused on ways of augmenting our everyday activities with help from pervasive computing, especially using multimodal or auditory interactive feedback. It is challenging to predict the type of technology that would be available in only a few years from now, whether the increase of computing power is going to slow down, or whether the miniaturization of computing artefacts will go on. It is possible that, in a near future, computing artefacts will “disappear” from our sensory world, their presence

being signaled only through their intended interaction. The evolution of technology will offer tremendous opportunities, and at the same time raises many questions and challenges, so much scientific as ethical ones.

Challenges and opportunities

The main challenge when working with cutting edge technology is its extremely rapid evolution leading to a quick obsolescence of the working platform. As an example, three generations of smartphones have been used in the time frame of this thesis. Because this evolution is driven by companies rather than academy, the development of functionalities often takes precedence over a comprehensive documentation. Moreover, security policies employed by the manufacturers are often restricting the freedom of the programmer. This may give rise to time-consuming difficulties when porting programs developed on a stationary computer onto a mobile platform. A common dilemma for a mobile application developer is to choose between spending time to port a project on a newer, less demanding platform, and maintaining a project on an obsolete and inconvenient system.

Recent handheld devices such as smartphones offer many new interaction possibilities, due to the increasing number of incorporated sensors: shaking, tilting, scratching, attaching to the body, squeezing are ways of controlling programs that are becoming more and more common. In addition, networking protocols can be integrated to these programs to create social interaction. On the one hand, local networks enable collaboration in a single location where results of the social interaction can be shared. In Section 4.2 we present a collaborative version of the program *Moodifier*, allowing participants to “jam” by controlling each a performance parameter. On the other hand, using global network offers the possibility for different participants to collaborate from different cities and countries anywhere in the world.

In our study of the current market of Android applications (Paper IV), we point out that few applications are adapted to their context of use, i.e. a mobile platform offering new possibilities of interaction. Designing applications in a creative way by exploiting a maximum of these opportunities represents a real challenge.

3.2 Tools for sound design and sound interaction

In a first step, the sonification model *Pure tone* presented in Section 4.2 was implemented in offline conditions using the classical programming language Python. It appeared very soon that this was not the appropriate tool for developing advanced sound synthesis models. On the other hand, we could efficiently manage interaction with different sensors in a rather straightforward way by using the program Python for S60, a port of Python for Symbian OS. For this reason, it was chosen to implement the KTH rule system for expressive music performance, the desired sound synthesis being limited to MIDI sounds.

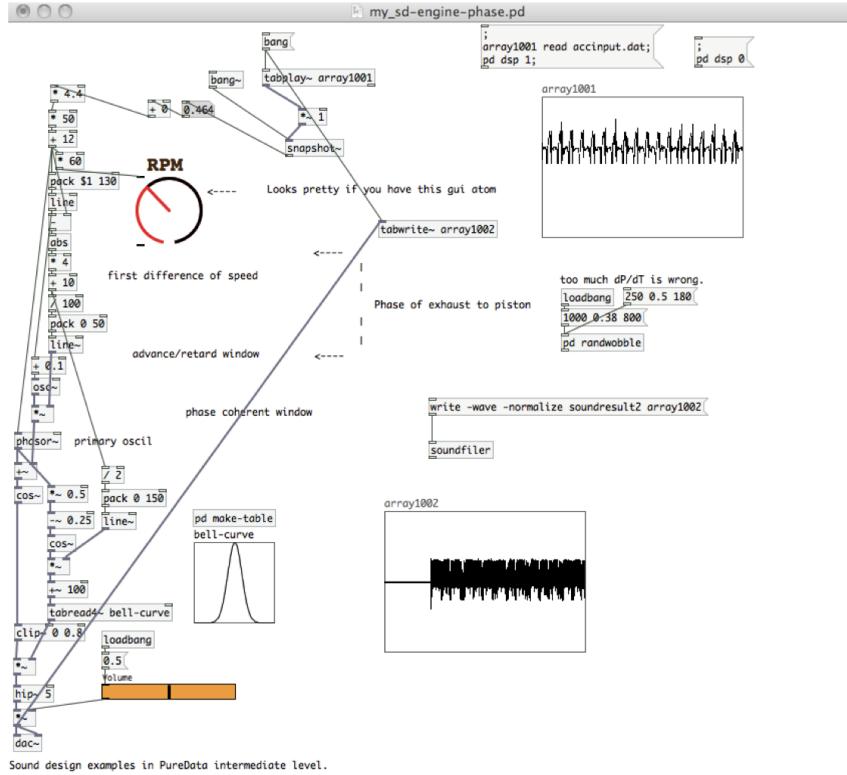


Figure 1: PureData patch used for creating sound samples corresponding to the model *Car engine*, taking rowing acceleration samples as input. The patch was adapted from Farnell.

When the need was felt for a more powerful synthesis tool, we began prototyping using the visual programming language PureData, specifically designed for real-time processing of data, and in particular audio data. PureData runs visually defined scripts called *patches* (see Fig. 1), where functions are represented by objects connected to each other to define a data flow. At the time we initiated the sonification design, to have such a powerful tool running on a commercial mobile platform remained in the realm of the unthinkable. Yet, in the course of just a few years, several ports of the PureData core distribution have been developed, among which the library libpd (Brinkmann *et al.*, 2011) allowing to embed the PureData sound engine into stand-alone applications on the Android platform.

More recently, significant efforts have been made by the PureData community to port the program on the Raspberry Pi (Raspberry Pi Foundation), a credit-card-sized computer that may represent the beginning of a new era for mobile devices — or should we say mobile computers?

Chapter 4

Results

4.1 Database of Android applications related to sound and music computing

In Paper IV, we investigated the different types of sound and music applications existing on the Android platform. The objective was to create a collaborative database showing the diversity of applications, and constituting a dynamic resource for developing novel applications. This resource, specially tailored for the Sound and Music Computing (SMC) community, would allow developers to discover niches for innovative applications. A preliminary database based on non-hierarchical keywords (called *tags*) was created in a non-collaborative environment and filled up with more than 1000 applications. Tags were assigned to approximately 150 of them in order to conduct a preliminary analysis. Ten broad categories of applications were defined based on the tags present in the database. The database was designed to be flexible, allowing for a dynamic evolution of the list of tags and of the broad categories. An implementation of a collaborative version of the database in the style of a wiki was realized by Guillaume Bellec, an undergraduate student, but has not been made available online yet.

4.2 Applications developed

ModifierLive

In Paper III, we presented an example of application based on the KTH rule system for expressive music performance. A simplified version of this system was ported to the mobile operative system Symbian, using the platform Python for S60 along with a C++ wrapper for real-time control of the smartphone's MIDI synthesizer. In this version a limited number of rules were implemented, based on combinations of four parameters of the performance: tempo, sound level, articulation, and phrasing. The mobile application, called *ModifierLive*, is able to change the expression of



Figure 2: The three modes of interaction implemented in *ModifierLive* were (a) a set of four sliders, controlled by the phone’s keyboard (b) a virtual ball navigating in the activity-valence space, controlled by tilting the phone (c) a virtual marble in a box controlled by expressive gestures (e.g. shaking, tilting).

the performance of a MIDI file in real-time. It includes three different modes of interaction, shown in Fig. 2. In the following we present briefly the program with focus on user interaction. A more detailed description can be found in Fabiani *et al.* (2011).

In the first mode, the user can control the value of each performance parameter independently by selecting and acting on sliders with the phone’s keyboard. This simple interaction mode was extended to a multi-user collaborative application, the architecture of which is shown in Fig. 3. In this collaborative mode, the music file is played by one phone having the role of a server, to which other phones can connect as clients via Bluetooth. Each client can then “book” a slider, modify the value of the corresponding parameter, and “release” it to allow another client to take control of the slider, or to book another slider. In this mode, several users can interact to create collaborative performances.

In the second interaction mode, the user controls a virtual ball navigating on the screen by tilting the phone. The screen represents the activity-valence space as described in Section 2.1, i.e., the coordinates of the virtual ball affect the performance in real-time. In addition, the size and the color of the ball vary depending on its position in the activity-valence space in order to reinforce the expressivity of the interface, since it is known that certain colors are naturally associated to emotional states (Bresin, 2005).

Finally, the third interaction mode emulates a box filled with marbles, and extracts features from the acceleration data when the user performs expressive gestures by, e.g., shaking or tilting the phone. Two different models for recognizing expressive gestures were used and are evaluated in Paper III. The first model was based on a decision tree elaborated after collecting data from expressive gestures

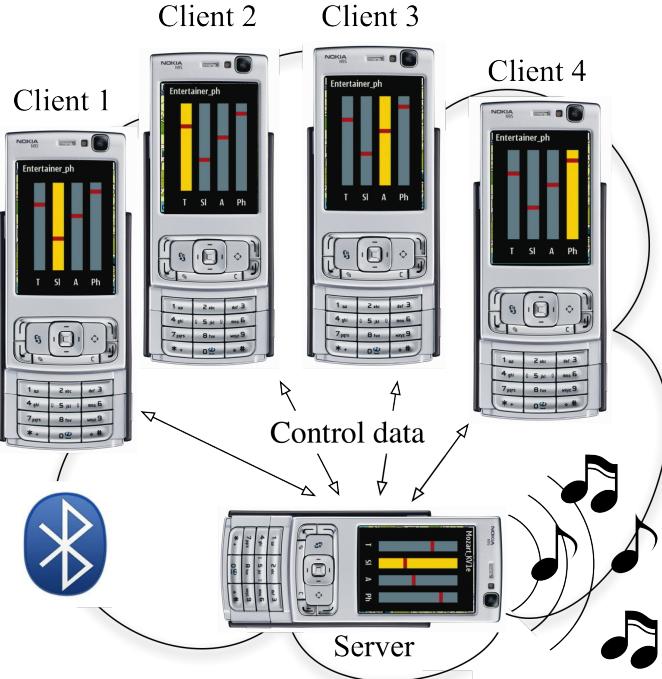


Figure 3: In the collaborative mode of *MoodifierLive*, a local network is created via a Bluetooth protocol. One smartphone plays the role of a server and receives connections from other phones (clients). The server plays an expressive performance of a music file according to the performance parameter values of four sliders. Each client can manipulate a slider to modify a given parameter (tempo, sound level, articulation, or phrasing), resulting in a collaborative performance.

with a handheld device. The model uses jerkiness and velocity of the movement to recognize a category of expressive gesture, classified according to four emotions (happy, angry, sad, tender). The program plays back the expressive performance according to the recognized emotion, which corresponds to a predefined set of performance parameter values. The second model corresponded to a continuous mapping of the acceleration components onto coordinates in the activity-valence space, and was implemented by the authors prior to the collection of acceleration data.

A demonstration of the three modes of interaction was presented in the framework of the 2009 Agora festival at IRCAM in Paris, France. For this occasion, the phone was used as a controller and the performance was synthesized by a Disklavier piano.

Sync'n'Mood

The application *Sync'n'Mood* also uses the KTH rule system for expressive music performance. It was designed to provide a concurrent feedback to two persons, each one holding a mobile phone in one hand, and trying to perform gestures synchronized with each other. Unlike *MoodifierLive*, *Sync'n'Mood* was not designed as a stand-alone mobile application. Instead, the mobile phone is used as a controller sending acceleration data to a stationary computer. On this computer, the software EyesWeb computes an index indicating the level of synchronization between the two controllers. This index is then communicated, together with the acceleration data, to the program *PerMORFer* which implements the KTH rule system for expressive music performance. *PerMORFer* performs the sonification of synchronization and acceleration via advanced high-level mappings described in details in Paper II and presented in Section 4.3.

Sync'n'Mood was evaluated and compared to two other sonification prototypes having the same purpose: to help two persons to synchronize their hand gestures when holding a mobile phone. This evaluation is detailed in Paper II and its results are summarized in Section 4.4.

Interactive sonification of rowing

Four models for interactive sonification of rowing were designed and evaluated via listening tests (Paper I) and on-water experiments (Paper VI). The main objective of these models was to provide an interactive auditory feedback related to the boat motion of a single sculler. By enhancing the training of elite athletes with interactive sonification, we aimed at altering their individual rowing technique in such a way that they would improve their performance. In order to create an interaction loop allowing the rowers to perceptualize the result of their actions in real-time, and therefore try to optimize the timing and magnitude of their gestures, we decided to provide a continuous feedback of kinematic quantities. These quantities (acceleration and velocity fluctuations) were electively displayed by the four models in the form of musical and environmental sounds.

Designing a continuous display is a challenging task, especially in the context of elite training where users are focused on a primary task that is physically demanding. Whereas such a display is more informative than a discrete sound feedback, it can be judged intrusive if poorly designed. Our first concern was therefore the acceptance of the system by elite rowers. An evaluation of the design of the four different models was conducted in offline conditions with both elite and casual rowers. Interactive experiments were then conducted with members of the Swedish national rowing team, and the impact of the sonification on the rowing technique was investigated.

The four models were implemented on a mobile platform, the whole setup being compact enough to be taken onboard. We used commercially available smartphones

to handle kinematic data and generate the interactive sound feedback, in order to develop applications that could be distributed to the public.

The first model, called *Musical instruments*, was implemented on a Nokia N95 mobile phone running Symbian OS, using the program Python for S60. In this model, the smartphone receives acceleration data from an external accelerometer via Bluetooth, and maps velocity fluctuations (computed on the fly) to the center frequency of a trill played by a musical instrument using the MIDI synthesizer. Pizzicato strings were used in the experiments, but this setting can be easily changed according to the preference of the user. A discrete sonification of the stroke rate was added using percussive sounds (drum and bell).

The three other models were implemented on a Samsung Galaxy SII using the program ScenePlayer, a port of PureData on Android OS. Acceleration data was read from the phone's accelerometer and used to synthesize the sound feedback. The model *Pure tone* mapped velocity fluctuations to the pitch of a pure tone. The model *Car engine* synthesized the sound of a car engine, whose brightness was coupled to acceleration of the boat. The model *Wind* mapped velocity fluctuations to the loudness of a wind sound. Finally, a model consisting of a superimposition of *Car engine* and *Wind* was implemented for the on-water experiments.

The two models *Car engine* and *Wind* use everyday metaphors in order to minimize the cognitive load on the athletes. The model *Wind* was the most appreciated, during both the listening tests and the interactive experiments. On the other hand, *Car engine* was judged annoying by a majority of participants, some of them stating that sounds of the nature were the best design choice.

4.3 Mapping data to sound

When designing sonification applications, the choice of the mappings to use is a fundamental issue. A mapping is an association between a data dimension and an auditory dimension. This association is characterized by a mapping function determining how a change in the input dimension should be reflected in the auditory dimension. A mapping has to be chosen carefully in order to make patterns, variations, and relationships in the data perceivable. If the domain and codomain of the mapping function can be ordered, a polarity of the mapping may be defined (if the mapping function is monotonous). To determine if a certain mapping is more informative than others, an evaluation has to be conducted, for example by performing psychoacoustical tests assessing how consistent a large number of people can be when perceiving effects of the mapping. Several such experiments have been conducted by Walker *et al.* (Walker, 2002; Walker and Kramer, 2005; Walker, 2007; Walker and Mauney, 2010), who investigated the perceptual effect of a few associations between conceptual data dimensions and auditory dimensions. The perceptual scale of a mapping could be derived by averaging the magnitude of numerical answers corresponding to the participants' assessment of the mapping.

Results of the systematic review

In order to get an overview of the use of mappings in sonification works, we conducted a systematic review sonification of physical quantities (Paper V). We analyzed a total of 179 scientific publications by identifying the type of sonic material and the set of mappings that were used in practical sonification applications. The first step was to define a set of conceptual dimensions in both the physical and the auditory domain. This was performed by grouping low-level, domain-specific dimensions that shared the same physical nature. These conceptual dimensions were gathered in high-level categories in order to perform a statistical analysis at different levels. The classification of the dimensions was performed in a flexible way that enables dynamic evolution and integration of special dimensions that can be divided into several classes or distinguished according to the scale used.

By counting the number of occurrences of each mapping, we could determine the most popular associations. Spatial dimensions of the sound were found to be used mainly for sonifying kinematic quantities. Pitch was the auditory dimension that was used the most, independently of the sonified physical dimension. Apart from associations involving pitch, the most popular mappings followed the logic of ecological perception, many sonification designers aiming at emulating a real physical phenomenon in their auditory display.

A publication database gathering a large number of sonification-related works was built up. Analyzing the year of publication of the entries of this database provides a historical overview of the research in sonification. From the publications included in this database and analyzed in the systematic review presented in Paper V, a mapping database was built. Several approaches were presented to analyze the contents of this mapping database. Mappings that had been evaluated in a rigorous way were given a particular attention, one of the goals of the study being to determine what mappings had been assessed as successful or unsuccessful. Our systematic review showed that only a small percentage of the mappings present in the database had been assessed, highlighting the need for a more systematic evaluation of sonification displays.

Sonification mappings used

The sonification applications developed in the framework of this thesis and presented in Section 4.2 make use of sonification mappings that were implemented before the results of the systematic review were available. We summarize them below, according to the classification defined in Paper V.

For sonifying rowing kinematics, we used the following mappings: in the models *Musical instruments* and *Pure tone*, *Velocity* of the boat was mapped to *Pitch*. In the model *Wind*, *Velocity* was mapped to *Loudness*. Finally, in the model *Car engine*, *Acceleration* was mapped to the *Brightness* of the synthesized sound.

In the program *Sync'n'Mood*, the synchronization of the gestures can be assimilated to the dimension *Phase* of our classification. It was mapped to *Melody*

lead. In addition, the **Energy** associated to the gestures was mapped to **Tempo**, **Loudness** and **Articulation** simultaneously, combined in a specific way to influence the **Performance activity level**.

In *MoodifierLive*, the first interaction mode (sliders) is not considered as sonification. In the second and third interaction modes, **Acceleration** data is used to control the expressive performance, via four parameters: **Tempo**, **Loudness**, **Articulation**, and phrasing (which does not correspond to a dimension of our classification). However, the models involved in this application are defined at a higher conceptual level than the dimensions defined in the systematic review, combining several of them in a meaningful way, both concerning the input physical dimensions (expressive gestures) and the auditory dimensions (expressive performance). **Performance activity level** is present in our classification, but not the other dimensions such as the valence or the different types of emotional performances.

4.4 Evaluation results

In each application developed, an evaluation of the design and effectivity has been conducted. In *MoodifierLive*, two models of recognition of expressive gestures were compared by participants of an experiment, who found the model based on the decision tree to correspond better to their movement than the model defined a priori by the authors.

The sonification model used in *Sync'n'Mood* was compared to other models of sonification of synchronization, both in terms of user preferences and actual performance. On average, participants preferred the control condition, i.e. when they were trying to synchronize their gestures without any sonification. This may reflect a lack of understanding of the purpose of the experiment, because few practical applications of the system could be discerned, or an appreciation that the task was not vital and therefore required no assistance. When analyzing the performance of the participants, it was shown that a learning effect occurred for all three sonification models, the participants becoming more effective towards the end of the experiment. However, the learning effect corresponding to *Sync'n'Mood* was the weakest of the three models. This is probably due to the excessive complexity of the mappings used, a problem that was mentioned by several participants in an evaluation questionnaire.

An advanced evaluation was conducted for the models developed for the sonification of rowing. Function and aesthetics of the design were assessed in a non-interactive experiment. Results showed that basic characteristics could be perceived from sonified acceleration samples, but not more advanced characteristics such as the gender and experience of the rower. Rowers were found to correlate the perceived functionality of the models to the aesthetics. A ranking of the models could be established with respect to overall preferences, although interpersonal differences suggested keeping a large palette of sound models available, for the users to choose according to their preference. An evaluation of the influence of sonification on the

efficiency of the rowing technique was performed after interactive experiments. Efficiency linked to energy lost due to velocity fluctuations was measured, for which no significant influence of the sonification could be shown. Sonification was found to have a certain influence on the stroke rate for some participants, however no general trend could be revealed. The measure of efficiency related to velocity fluctuations was found to depend linearly on the stroke rate value, the rowing technique becoming less efficient as the stroke rate increases.

Chapter 5

Discussion

5.1 Contributions

In this section we detail the contributions of each of the papers included in the thesis.

Paper I

In this article we present the principle of sonification and motivate its use in the context of sport training. Based on acceleration data samples collected with athletes from the Swedish national rowing team, listening tests were conducted with both casual and elite rowers. Results show the potential of the sonification models to communicate basic characteristics of the data, while it is hypothesized that more advanced characteristics (e.g. gender, experience) would require the combination of several types of data (e.g. kinematics and kinetics). The evaluation conducted revealed marked interpersonal differences, some rowers stating that they could use any of the models except one, which was then chosen by other rowers as the only model that they would use. Acceptance rate of the sonification was rather low in general as compared to prior studies (Schaffert et al., 2009), and it was hypothesized that it would increase if the rowers could test the system in on-water experiments. Since the detailed acceptance rate of each model revealed differences between casual and elite rowers, it was suggested that elite rowers were more attracted towards models making use of environmental sounds in metaphorical associations, while casual rowers preferred models making use of musical sounds. A ranking could be established with respect to overall preferences and aesthetic qualities. The favorite model was the same in both cases, and consisted in a wind sound varying in loudness.

Paper II

This study investigated three interactive sonification models for the synchronization of two participants performing gestures with a handheld controller. A synchroniza-

tion index was computed in real-time reflecting the performance of the participants. The first auditory display (*Sync'n'Move*) used a multi-track recording to map the value of the synchronization index to the number of instruments rendered, rewarding participants achieving a good level of synchronization over a given period with the complete orchestration of the musical piece. The second model (*Sync'n'Moog*) mapped the synchronization index to the center frequency of a Moog filter applied to the same piece of music, making the auditory feedback less pleasant when the synchronization was worse. The third model (*Sync'n'Mood*) used the KTH rule system for music performance, mapping the overall energy of the gestures to the activity of the rendered performance, as well as mapping the synchronization index to a time offset of the melody in relation to the accompaniment. An evaluation of the performance was conducted for the three different models, with and without visual contact between the two participants involved, and in a control condition (without sonification). According to questionnaires answered by the participants, the control condition was preferred and perceived as inducing the best performance. However, an analysis of the synchronization index indicated that the sonification helped the participants to synchronize, especially in the last part of the trials, suggesting that a learning effect occurred. No significant difference was observed between the three sonification models with respect to preferences and perceived performance. The analysis of the synchronization index indicated that *Sync'n'Mood* was outperformed by the two other models. By choosing a particular type of sonic material as a basis for the auditory display, namely a piece of music unfolding in time, our models appear to be outside the borders of the most restrictive definitions of sonification (e.g. Hermann, 2008). Arguing that, according to its original definition, such models using high-level mappings should be considered unambiguously as sonification, we point out the need for a broadened definition.

Paper III

This paper introduces *MoodifierLive*, a mobile phone application using the KTH rule system for expressive music performance to perform the sonification of expressive gestures. Experiments were carried out to compare two different models mapping gesture data to expressive performance, in order to determine which one led to an auditory feedback perceived as more consistent with the gestures of the participants. The authors designed the first model *a priori*, whereas the second one was based on expressive gesture data collected in a prior experiment. In both cases, features of acceleration data were extracted and used to match a corresponding emotion via a classification tree. Four emotions were recognized (anger, happiness, sadness, tenderness), and corresponded to specific coordinated in the activity-valence space of the performance rendered by the program. An evaluation based on individual questionnaires was conducted, and showed that participants judged the auditory display based on previously collected gesture data corresponded better to their own gestures than the model defined *a priori*.

Paper IV

In this short study, we examined the state of the market of Android applications related to Sound and Music Computing (SMC). By selecting a large number of applications from the official, centralized application database (i.e. *Google Play*), we created a database based on non-hierarchical keywords (tags) and centered on the needs of the SMC research community. From the database we could derive a classification of the existing applications into ten broad categories. A preliminary analysis of the repartition of the applications among these categories suggested that the great majority of them are dedicated to extremely simple tasks such as sound effects, classical multimedia players and sound recorders. Many applications of the database were found to consist in a simple port of a program already existing on a stationary platform. On the other hand, rather few applications made full use of the opportunities enabled by a mobile platform. Finally, it was shown that only a marginal proportion of applications were dedicated to sound programming engines, which could therefore direct the efforts of future application development projects by the SMC community.

Paper V

This systematic review is focused on more theoretical aspects of sonification. The main objective of the work is to organize the knowledge accumulated in nearly 20 years of research on auditory display, via an analysis of the mappings used in sonification applications throughout history. In this way, we could learn from previous research which associations between physical and auditory dimensions have been the most popular, successful or unsuccessful, and build a foundation for future sonification works. In the introduction of the paper, we present an overview of the field of sonification, starting from the definition, describing the different sonification techniques commonly used, and taking a historical perspective to introduce examples of sonification applications (among which the first known use of an auditory feedback in a Geiger counter). We define a sonification mapping as a function defined from a data domain space (here we restrict the study to physical data) to the auditory domain. We set out the framework of the systematic review by defining and describing the classification spaces for physical dimensions (domain of the mappings) and for auditory dimensions (codomain of the mappings). The studies considered in the systematic review, representing 179 scientific publications gathered in 60 research projects, are then presented in a table along with the sonic material and the list of mappings used in the sonification application. A statistical analysis is performed on the frequency of mappings identified in the considered publications: a simple census of the most popular mappings is presented, the use of auditory dimensions is investigated regardless of the input physical dimension, high-level trends (corresponding to broad categories of dimensions) and intermediate-level trends (corresponding to generic dimensions, more suitable in the scope of proposing design guidelines) are analyzed. In addition, one example

of multi-class dimension and one of multi-scale dimension are given and used in the analysis, tackling the need for a multi-level structure as well as illustrating the plasticity of our classification. Different labels were assigned to mappings that were assessed rigorously via psychophysical experiments and found either successful or unsuccessful, as well as mappings not implemented but mentioned as potential developments of a sonification system. The small number of assessed mappings enabled only elementary statistical analyses to be performed, which illustrates the need for a more systematic evaluation of sonification works. We also present alternative ways of analyzing the mapping database, such as a keyword-based analysis of low-level dimensions, and the study of project-related trends. The publication database is used to provide other types of information such as the historical distribution of sonification works, a classification of the projects in terms of primary function, as well as the type of sonic material constituting the different auditory displays. Finally, a new approach based on mappings is proposed in order to clarify the definition of the concept of sonification.

Paper VI

This paper is a follow-up study of Paper I, consisting in an evaluation of sonification models for rowing in an interactive context. A similar evaluation was conducted in order to assess the preferences of the rowers, this time after having tested the system interactively. As it had been hypothesized, the acceptance rate was in average higher than in non-interactive conditions. The wind model was, once again, the favorite model of the rowers. Aesthetics of another model, namely a car engine sound varying in brightness with acceleration, was harshly criticized by most of the participants. Nevertheless, it could be noticed that superimposing this least preferred model to the most preferred model led to a significantly increased acceptance rate. The efficiency of the individual rowing style was also assessed in order to observe whether some improvement could be observed when rowing with help from an interactive sonification system. To this end, we defined a measure of velocity fluctuations, which are known to be detrimental to rowing efficiency. This measure was applied to average cycles normalized in time for each participant. No effect of sonification could be observed on this measure. Stroke rate regularity was observed when using a discrete and a continuous auditory display. Although the presence of the auditory feedback had an effect on the stroke rate value for some participants, no general trend could be observed. Interestingly, the measure of velocity fluctuations was found to depend linearly on the stroke rate value, indicating a decrease of efficiency at a higher stroke rate.

5.2 Conclusions

Developing applications for the interactive sonification of body motion incorporates many facets, from practical issues to theoretical considerations. Throughout this

doctoral project we have focused on (1) hardware as well as software issues for implementing robust applications on a mobile platform, (2) design aspects of sonification models — both from a functional point of view and with respect to aesthetics, (3) evaluation of sonification systems, and (4) evaluation of crucial components of these systems such as mappings.

A significant part of our work is in the field of sonification theory. In an attempt to provide an overview of the field of sonification, we conducted a systematic review of mapping strategies for the sonification of physical quantities in Paper V. In the context of that paper, we presented the current status of research and uses, we set them in a historical perspective, and we proposed a new approach for defining the boundaries of sonification. Our systematic review highlights the need for a more systematic evaluation of sonification mappings. The results of this study, showing the most popular mappings, cannot be considered as design guidelines as long as no proper mapping evaluation has been conducted. Therefore, we suggest assessing these popular mappings in priority.

Practical applications of interactive sonification have been developed in different contexts. In the case of elite sports training, the acceptance of the system was one of our primary concerns. The example of elite rowing has provided an illustration of the needs for multimodal feedbacks to be implemented on a mobile platform. Assessment of the design showed general trends in the aesthetic value given to the different models, while still reflecting individual preferences in the way athletes experienced the sonification. Evaluation of the effect of the sonification system on kinematic quantities revealed no significant influence of the system, independently from the model tested. That is, interactive sonification did not influence the performance after a first experiment. On the other hand, no negative effect was shown, which suggests that the system was well accepted and could be used over longer periods for investigating a potential learning effect.

Interactive auditory feedback based on expressive music performance allowed us to design high-level sonification mappings of expressive user gestures when interacting with handheld devices. We developed advanced interaction modes by taking advantage of the opportunities offered by progress in mobile technology, i.e. the increase of computational power and the availability of embedded sensors.

The continuous evolution of mobile technology observed during the course of this work confirms the importance of focusing research on the field of user interaction and perception for the development of robust future applications that can answer in a natural and intuitive way to human actions.

Bibliography

- Roberto Bresin. 2005. What is the color of that music performance? In *Proceedings of the 31st International Computer Music Conference (ICMC 2005)*, pages 367–370, Barcelona, Spain.
- Peter Brinkmann, Peter Kirn, Richard Lawler, Chris McCormick, Martin Roth, and Hans-Christoph Steiner. 2011. Embedding Pure Data with libpd. In *Proceedings of the 4th International Pure Data Convention (PdCon 2011)*, Weimar – Berlin, Germany.
- Giovanni Buccino, Ana Solodkin, and Steven Lawrence Small. 2006. Functions of the mirror neuron system: implications for neurorehabilitation. *Cognitive and Behavioral Neurology*, 19(1):55–63.
- Baptiste Caramiaux, Patrick Susini, Tommaso Bianco, Frédéric Bevilacqua, Olivier Houix, Norbert Schnell, and Nicolas Misdariis. 2011. Gestural embodiment of environmental sounds: an experimental study. In *Proceedings of the 11th International Conference on New Interfaces for Musical Expression (NIME'11)*, pages 144–148, Oslo, Norway.
- Marco Fabiani. 2011. *Interactive computer-aided expressive music performance*. PhD thesis, KTH Royal Institute of Technology, Stockholm, Sweden.
- Marco Fabiani, Gaël Dubus, and Roberto Bresin. 2011. ModifierLive: interactive and collaborative music performance on mobile devices. In *Proceedings of the 11th International Conference on New Interfaces for Musical Expression (NIME'11)*, pages 116–119, Oslo, Norway.
- Andy Farnell. URL <http://obiwannabe.co.uk>. Personal website. Accessed August 30, 2013.
- Neville Horner Fletcher. 2004. A simple frequency-scaling rule for animal communication. *Journal of the Acoustical Society of America*, 115(5):2334–2338.
- Anders Friberg. 2006. pDM: an expressive sequencer with real-time control of the KTH music-performance rules. *Computer Music Journal*, 30(1):37–48.

- Anders Friberg, Roberto Bresin, and Johan Sundberg. 2006. Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology*, 2(2–3): 145–161.
- Bruno Lucio Giordano and Stephen McAdams. 2006. Material identification of real impact sounds: effects of size variation in steel, glass, wood, and plexiglass plates. *Journal of the Acoustical Society of America*, 119(2):1171–1181.
- Andrew Godbout and Jeffrey Edwin Boyd. 2010. Corrective sonic feedback for speed skating: a case study. In *Proceedings of the 16th International Conference on Auditory Display (ICAD 2010)*, pages 23–30, Washington, DC, USA.
- Thomas Hermann. 2008. Taxonomy and definitions for sonification and auditory display. In *Proceedings of the 14th International Conference on Auditory Display (ICAD 2008)*, Paris, France. CD-ROM.
- Thomas Hermann, Andy Hunt, and John G. Neuhoff, editors. 2011. *The sonification handbook*. Logos Publishing House, Berlin, Germany.
- Andy Hunt, Thomas Hermann, and Sandra Paulette. 2004. Interacting with sonification systems: closing the loop. In *Proceedings of the 8th International Conference on Information Visualisation (IV'04)*, pages 879–884, London, UK.
- Evelyne Kohler, Christian Keysers, Maria Alessandra Umiltà, Leonardo Fogassi, Vittorio Gallese, and Giacomo Rizzolatti. 2002. Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582):846–848.
- Stephen Michael Kosslyn, Giorgio Ganis, and William Ladd Thompson. 2001. Neural foundations of imagery. *Nature Reviews Neuroscience*, 2(9):635–642.
- PureData. URL <http://puredata.info>. Accessed August 30, 2013.
- Raspberry Pi Foundation. URL <http://www.raspberrypi.org>. Accessed August 30, 2013.
- Matthias Rath. 2003. An expressive real-time sound model of rolling. In *Proceedings of the 6th International Conference on Digital Audio Effects (DAFx-03)*, pages 165–168, London, UK.
- Bruno Hermann Repp and Amandine Penel. 2004. Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*, 68(4): 252–270.
- Davide Rocchesso and Roberto Bresin. 2007. Emerging sounds for disappearing computers. In Norbert Streitz, Achilles Kameas, and Irene Mavrommatti, editors, *The Disappearing Computer*, volume 4500 of *Lecture Notes in Computer Science*, pages 233–254. Springer Berlin Heidelberg.

- Davide Rocchesso, Roberto Bresin, and Mikael Fernström. 2003. Sounding objects. *IEEE MultiMedia*, 10(2):42–52.
- Carla Scaletti. 1994. *Auditory display: sonification, audification and auditory interfaces*, chapter 8: Sound synthesis algorithms for auditory data representations, pages 223–251. Addison Wesley Publishing Company.
- Claes von Hofsten. 2009. Action, the foundation for cognitive development. *Scandinavian Journal of Psychology*, 50(6):617–623.
- Bruce N. Walker. 2002. Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of Experimental Psychology: Applied*, 8(4):211–221.
- Bruce N. Walker. 2007. Consistency of magnitude estimations with conceptual data dimensions used for sonification. *Applied Cognitive Psychology*, 21(5):579–599.
- Bruce N. Walker and Gregory Kramer. 2005. Mappings and metaphors in auditory displays: an experimental assessment. *ACM Transactions on Applied Perception*, 2(4):407–412.
- Bruce N. Walker and Lisa M. Mauney. 2010. Universal design of auditory graphs: a comparison of sonification mappings for visually impaired and sighted listeners. *ACM Transactions on Accessible Computing*, 2(3):12:1–12:16.
- Bruce N. Walker and Michael A. Nees. 2011. *The sonification handbook*, chapter 2: Theory of sonification, pages 9–40. Logos Publishing House.

Part II

Included papers

