Extending the control of a musical tangible tabletop with mobile phones

Alexandros Katsaprakakis

Master Thesis UPF / 2011

Thesis Director:

Dr. Sergi Jordà

UNIVERSITAT
POMPEU FABRA

Abstract. This project proposes the use of the sensory features of modern mobile phones that were previously unavailable for the wide consumer public, such as the accelerometers and the multi-touch surfaces, as an input for a musical tangible tabletop. The development of a mobile application operating simultaneously with the reactable is carried out based on a conventional wireless connection. The application is a formant voice synthesizer controlled by both the mobile device and the reactable. The proposed system seeks to give an additional degree of freedom in the musical experience and expression with this instrument and tries to take advantage of the advancing mobile technology to extend the control of the reactable and provide alternative ways of musical interaction.

Keywords: mobile music, tangible tabletops, mobile phones, accelerometers, multitouch, reactable, shared control

Table of Contents

Acknowledgements. I would like to thank the ones that have been beside me during this tough but also exciting period. They have supported me, inspired me and have tolerated my anxiety. This project would have been impossible without the help, the ideas and the directions of Sergi Jordà, my supervisor and teacher. Thanks to Carles Julià, that patiently helped me set up my experiments. I would also like to acknowledge the helpful suggestions of the MTG members Jordi Bonada and Mathieu Bosi. Last but not least, the members of "Our little choir": John O'Connell, a wonderful improvising 'maestro', Srikanth Cherla, Jorge García and Stelios Togias, who has helped me out further in testing my application. I hope I'll get the chance to 'sing' with them again..

It was musically quite an interesting period studying and experimenting with you all!

1. Introduction

Mobile phone technology has experienced a tremendous growth over the last few years. The use of 'smart' mobiles is now widespread, while ten years ago fixed phone devices used to be the standard. Nowadays, a serious percentage of the global population owns a mobile phone and uses it in a number of ways, from the conventional telephone calls to a wide range of diverse and complex applications. A lot of people now use their mobile phone for entertainment, while there are many that need it in their professional life.

In addition, the computational power and the sensory capabilities of modern mobile phones are increasing, blurring the distinction between phones and laptops. High computational capacities along with multiple and powerful sensors are already embedded in small devices, pretty easy to handle and carry around with. It comes therefore as no surprise that an interest in using such systems as musical instruments is observed. Out of the whole amount of mobile applications, a percentage of around 10% is reported to be related to music at the time being. Some of them come out even a few days after the release of some mobile device. Furthermore, the latest sensory additions to these devices, such as the accelerometers and the multitouch screens are means of interaction previously unavailable to the consumer public with the standard computers and laptops.

On the other hand, the use of tangible tabletops, as well as the research activity around them is constantly increasing, aiming to exploit their public and collaborative features. Music creation stands out as a remarkable field for research as well as for product development. The complex nature of musical interaction and collaboration seems to somehow fit with the 'decentralized' and collective interactive scenario that a tabletop implies. Systems seeking new ways of collaborative music creation and performance based on a tangible tabletop have already emerged. The reactable [1], a multi-user electro-acoustic musical instrument with a tangible tabletop user interface has an already remarkable life of about six years. Designed at the Music Technology Group of the Universitat Pompeu Fabra, it was planned since the beginning as an instrument focusing on collaboration, with no leading position nor fixed roles. It is nevertheless a promising platform for integrating with advanced mobiles. However, an integrated scheme of tabletops and mobiles has not been exhaustively researched yet.

The main goal of this project is to find suitable ways of using the sensors of mobile phones that are not available in desktops and tabletops, extending in this way the input capabilities and the control of the reactable. Till now, the musician playing the instrument can interact with it by placing different objects on its surface and controlling their relative position, angle and connections between the objects or by directly touching the display with one finger. Ongoing research tries to extend these capabilities by using force sensors on the perimeter of the instrument to detect pressure and percussive events [2]. A different part of current research is dealing with the extension of the interaction to the 3D space on the table's surface, above and around it, while another direction treats collaboration under the scope of implicit and explicit interaction and control. However, the objects used remain passive and the actual interaction between them and the tabletop is based on reacTIVision, a computer vision mechanism [3] that tracks markers printed on the surfaces of the objects (Fig. 1.1). This study introduces the use of active objects such as the widely available mobiles as a way of interaction with the musical instrument. Additionally, the wireless connection capabilities that mobiles offer is a step towards extending this interaction between the performer and the instrument away from the surface of the table, or even away from the actual place where the tangible tabletop is. The mobile phones are chosen for their availability and relative low cost. By assuming that a large percentage of the audience will carry such devices in a performance context, a lot of possibilities open up to interact with them in some way, from giving some visual feedback of the performance on their private displays to invite them to actually join the performance.

Figure 1.1: Four markers from the reacTIVision 'amoeba' set

However, a specific scenario had to be set for the study to be feasible and able to deepen in musically interesting issues, such as expressiveness, distributed or shared music control and the 'connectedness' of a performer using the mobile and the one using the reactable. A realistic one is to create an object that when placed on the surface of the reactable will establish a wireless connection with a preselected mobile device by using a specific marker like the ones mentioned. Let's call it the 'mobile' object. The mobile device is equipped with the required software and the application under development and acts as both an input and a block of the synthesizing process. Further connections, processing or effects happen on the surface of the reactable, just as with a normal existing object. This mobile object helps avoiding an in-depth transformation of the architecture of the instrument and enables us to focus more on the application and the interaction design and development. On the other hand, such approach does not limit the collaborative possibilities (e.g. connecting many mobiles at the same time), instead provides an easy way to control and visualize such connections on the surface of the tabletop by using multiple 'mobile' objects.

An important design goal is certainly to give the feeling to both performers and audience that the mobile instruments or applications are actually connected to the reactable and at the same time being perceived as individual controllers on the hands of performers. The moves required by a performer to use the interface, whether small moves with the fingers or moves with the whole arms and body will also be critical to its perception as a musical interface for both performers and listeners.

2. Mobile music creation

Mobile phones have attracted the interest of musicians and artists since the very beginning of their commercial explosion. By that time one was just able to play different monophonic tunes -ringtones-, he could play the same tunes though wherever, by just carrying around a mobile. This mobility was the main attraction in most concerts or art installations where these ringtones took part[1].

Nowadays mobile music can serve as a term to refer to any musical activity using portable devices, from song sharing to complex musical interactions. Academically though it has been used to characterize systems that aim to exploit the portability of these devices to make music. For example, musical instruments have been developed using mobiles with place and space as their musical inputs in a scheme without the need of eye to eye contact [4]. The location of the mobile user -whether outdoor or indoor- seems to be quite an interesting musical input. Another good example is the musical instruments dealing with long distance collaboration, where social aspects emerge, raising issues of agency[5]. Remote collaboration generally appears as a core issue with the fast expansion of social networks and their adaption to the mobile platforms lately, a progress quite promising for digital musical systems in general. The International Workshop of Mobile Music Technology was established and held for four times in this direction but has been inactive since 2007. On the other hand mobile music starts playing a more and more important role in the NIME (New Interfaces for Musical Expression) conference.

Figure 2.1: Mobile phones playing just some monophonic ringtones in art installations

The work carried out in this project is about mobile phones as an alternative or extended way of controlling the reactable by a performer standing away from its surface. This short distance mobility and the benefits and limits from the integration with the fixed tangible tabletop compared to the performance with just the reactable are to be the centre of attention. The mobile phones were chosen to experiment with the extension of the tabletop's control mainly for their availability and relative low cost in contrast to custom-made sensory technology that can be cheaper but is one of a kind. Additionally, the rapid growth of the field constantly offers new ways of interaction previously unavailable, whether in a real-time performance context (that is our main interest), or by using the social and mobile features that seem to be for the moment in constant evolution.

---

[1] http://www.flong.com/texts/lists/mobile_phone/

a) Brief history and modern problems

The first mobile calls were demonstrated in mid 1940s, using devices that had such a weight and size that had to be carried around inside a car. Their weight and size were too much for them to be considered as handheld. The first “handheld” mobile call is attributed to Martin Cooper from Motorola, who in 1973 is said to have called his competitor in the race for the first 'real' cellular telephone, Joel Engel from Bell Labs. This very first mobile call thus seems to be quite indicative of the wild competition between companies that has existed since the very beginning (or even before!) and continuously sped up till today. This in turn, has led to an extremely fast technological development, that undoubtedly has given the mobile users a lot of potential. Nevertheless, it has also created a great deal of incompatibility issues between devices, operating systems, functions and file types that they support. Obviously, the same stands for the sensing capabilities of the mobiles. The various manufacturing companies, still in the competition race, favour one type of technology over the other, constantly trying to outride rivals in what has to be the next technological breakthrough. One easily finds out that apart from making available more and more sensors, this competition also results in vast incompatibility issues, different technologies employed, and what is more important, low amount of information and technology transfer, in-depth specifications and reliable accuracy measures.

Figure 2.2: The phone used by M.Cooper for his 'first' call and its size compared to later devices.

Users and developers alike are caught in the middle of all this. The first ones usually have to upgrade devices and software without having the actual need or full knowledge of the functionality and potential of the devices they already own, while mobile developers have now more complicated distribution mechanisms and more incompatible development platforms and tools to struggle with.

On the other hand, PDAs (Personal Digital Assistants) appeared in the early 90s, also known as palmtop computers, personal data assistants or personal information managers. In mid 90s, the first mobile phones with full PDA functionality were presented and today both terms are actually merged into 'smart-phones'. This term however does not seem to be particularly welcome yet by the majority of mobile users, maybe it is not so smart as a term anyway.

By the time of this writing the mobile market is governed by three or four major brands/cooperations: Google, Nokia, Apple, Blackberry plus a cross-platform corporation formed by Intel and Asus, Acer, Dell, Samsung, Fujitsu, Siemens that is beginning to emerge. Each platform promotes its own operating system, apart from some collaborations that are constantly negotiated and change fast. In terms of mobile applications store revenue though, the Apple App Store is a long way ahead from the online stores of other brands.

b) Sensors embedded

In an attempt to review the sensing capabilities of modern mobile phones that constitute potentially interesting musical controllers, it can be easily seen that different devices have a quite different setup. Nevertheless, a resume of the sensory features of a 'generic' setup is possible, having in mind that it is going to differ quite a lot in certain implementations:

- **Microphone**

  Usually there is one per device and it is optimized for close range recording. For music recording however it does not seem to be a reliable solution. External microphones are already on the market, promising better recordings.

- **Cameras**

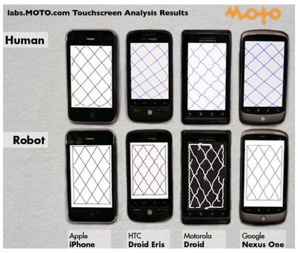
  Normally mobile phones came up with one camera. Nowadays, some models are equipped with two cameras, one for video call at the front side of the phone and one at the back used to record video and take photos. The back camera is accompanied by a flash. Lately, some companies have announced the integration of dual-lens cameras for 3D recording.

- **Electromagnetic antennas**

  The various mobiles base their functionality in GSM/3G antennas and networks. However, they also include a variety of other antennas that provide alternatives for communication, different from the basic function as a standard telephone. GPS, WLAN, Bluetooth and FM radio antennas are already mounted in almost all new models. DVB-H (Digital Video Broadcasting – Handheld) is not at the time a standard, but it is quite possible that it will be in a while. What would be an interesting musical input is location, whether outdoor or indoor. Indoor location can be estimated using WLAN neighbour signals. Outdoor location is obtained with the GPS or less precisely with the mobile cell tower identifier.

- **Touch screens, keyboards**

  Generally, mobile phones use capacitive sensing touch screens, but their implementations and accuracies differ a lot. Multitouch is the ability of the screen to recognize and receive data simultaneously from more than one finger. They are gaining territory in being the main form of interaction in smart-phones, however many users still complain of the lack of a hard keyboard in their models, as text entry on the soft keyboard and in a such confined space does not result very natural. Until now, this technology was available to standard desktop and laptop users in the form of external pads only.



- **Accelerometers and gyroscopes**

  They constitute a way of control quite new for the consumer public. Initially used to calculate orientation and adjusting the screen, they have attracted a lot of interest and now many applications have implemented some kind of control based on these sensing modes. Gesture control is made available by measuring the changes in rotation and displacement.

  The accelerometer is a device that measures acceleration in three axes, including gravity. The gyroscope measures orientation and rotation regardless of gravity. A combination of the two is available in modern smart-phones, allowing more accurate and stable recognition of movements in the 3D space than its preceding setup with just the 3-axes accelerometer measures.

- **Magnetometer**

Measures the strength and direction of a magneto-static field. It is used in conjunction with software to implement a digital compass.

- **Light brightness sensor**

Measures the amount of ambient light, primarily to adjust the display brightness. At the moment the access available to the iPhone's sensor is through a boolean variable.

- **Fingerprint recognition sensor**

Implemented in few models targeting professional use.

- **Private sensors**

Many sensors more are embedded in the interior of the devices. They are used for the needs of a normal functionality and they are not accessible by mobile developers. Examples include temperature and proximity sensors.

- **Environmental sensors next?**

c) Mobiles as instruments

The extreme expansion of mobile phones, along with their fast increasing capabilities, have attracted not only the interest of a widening academic circle, but also that of a large community of developers, hackers and simple users. A huge amount of music applications for mobile phones is already available (around 10% of the total amount of apps -50.000 out of 500.000!- available by the time of the writing), from simple accessories like tuners and metronomes to complex instruments and digital audio workstations. The vast majority though seems to be simulations of existing instruments and tools, or at least adaptions to the mobile platform. The easy -but also restricted in some cases- access to development kits, APIs and distributing mechanisms of the various mobile platforms has given rise to many free and low-cost applications. What is even more important is that it has permitted to experiment and risk much more when developing an interface, something impossible a few years ago in the design of commercial applications. Even new ways of selling music are now being commercially researched and start to appear.

It is important to mention that commercial apps emerge and evolve quite faster than the academic research, thus new technology or modes of functionality may be first implemented in commercial applications. However, there is still some tendency for the majority of products to imitate or at least stem from a traditional or digital existing instrument. This way, users have an idea of how the instrument is played and do not have to learn everything from scratch. However the embodied nature -the standalone quality- and the specific affordances of the various mobile devices may allow different, idiosyncratic mappings as examined in [6].

Although it is impossible to keep track of every commercial application, or even make a general review, some of them considered successful and popular are mentioned here. NaNoStudio[2] is a virtual recording studio app, MoogFiltatron[3] a realtime audio filter and effects

---

[2] http://www.blipinteractive.co.uk/index.php

[3] http://www.moogmusic.com/products/apps/filtatron

engine, while ReBirth[4] and Bleep!Box[5] are synthesizers including drum machines and sequencers. MusicStudio[6] is a complete production environment and TonePad[7] is a Tenori-On. All the above are considered quite successful at the time being and they are all imitations of former instruments. More abstract interfaces that are maybe less known include Jasuto Modular synthesizer[8], and Bloom[9]. They are all available at the iPhone app store as well as the online stores of some other platforms.

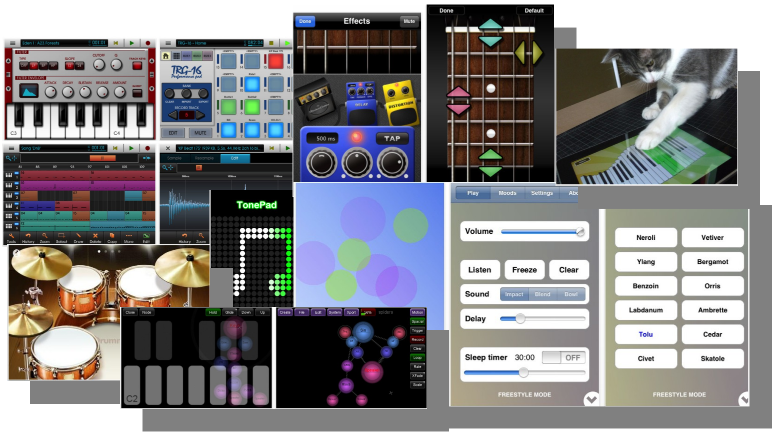


Figure 2.3: A small collage of commercial mobile apps

The reactable has also come up with a mobile version, for the iPhone, iPad and Android platforms [7]. The application is a virtual reactable controlled by fingers instead of objects and uses finger gestures to rotate the virtual objects and zoom in and out. The main goal of the design was not to exploit the mobile's interaction capabilities as in this study, but instead take advantage of the portability and bring the reactable closer to consumer´s reach. The cost of the application is extremely lower than that of a real reactable and its size is extremely smaller (Fig 1.2). The control of the instrument however is much more restricted. On the other hand, the potential of sharing performances over the various social networks seems to provide a quite

Figure 2.4: the iPad and iPhone reactable mobile on top of the surface of a real reactable

---

[4] http://rebirthapp.com/

[5] http://www.bleepboxapp.com/

[6] http://www.xewton.com/musicstudio/overview/

[7] http://www.tonepadapp.com/

[8] http://www.jasuto.com/site/

[9] http://www.generativemusic.com/bloom.html

interesting extension of the instrument as it has formed a live community of users around the instrument, that can now share performances, knowledge and feedback.

Another instrument that implements social networking features is Ocarina [8], an instrument that stems from the traditional flute and uses the microphone, accelerometer and the touch input of the phone. Breath controls intensity, fingerings on the screen define pitch and the accelerometer the amount and depth of the vibrato. In addition, GPS and WLAN are used together with a centralized network for users to exchange songs and information together with geographical data.

As for the abstract interfaces, it seems somehow more possible that they can intuitively exploit the sensory capabilities of the mobiles, as they are not built upon constraints of former technology. On the other hand, users usually have to spend more time and effort to learn how to play them. Physical based metaphors can be an intuitive aid for users, as they are explored in SoundBounce [9], where a virtual ball bounces on the mobile's surface and can be thrown from one performer to another. More complex metaphors have also been used, for example in [32], where the movement, spin and collisions of particles generate sound controlled by the multi-touch screen.

Generally, all apps tend to use multi-touch input as the basic form of interaction, as well as accelerometers to provide more delicate control features. However, the main control in almost all of them is accomplished using just the first finger. Other sensing capabilities, such as cameras, have been investigated as possibly interesting musical inputs. In Camus2 [10] an optical movement detection algorithm or alternatively a grid of virtual markers is implemented to compute spatial displacement among adjacent frames. With a frame rate of 15 fps, an amount of overlap between frames is assumed, at least for movements that are not too fast. This overlap is translated in a measurement of movement. The technique seems now even more promising with the recent developments in range imaging, enabling the acquisition of depth for a specific amount of points in a 2D image being captured.

Impact force and pressure based input for mobile music instruments have been investigated [11], but without extremely promising results yet. Quite recently, the MagiMusic project [12] proposed the use of the compass embedded in some phones, to achieve touch-less interaction in the 3D space around the device, towards more flexible and natural gesture based control. The compass is deviated by a simple magnet ring that someone can quite naturally wear in a finger.

Figure 2.5: the MagiMusic proposal for touchless interaction in a virtual guitar application

The technique is an extremely low cost solution and promises to enhance interaction with mobile devices.

Lately, the upgraded computational power of mobiles has also given the researchers the possibility to experiment with more complicated and computationally expensive tasks. Cognitive architecture and machine learning techniques are explored in [33], as a first step of exploring the possibilities of such an integration.

d) Development tools

Apart from instrument applications, some toolkits and generic instruments have also emerged, facilitating the development of musical applications. MoMu[13] is a toolkit developed at Stanford intended to act as an abstraction layer on top of the iPhone Software Development Kit. It provides easy access to the sensory data and simple utilities useful for mobile music creation. Another environment with the same goal, but a different perspective is urMus[14]. It is designed to assist musical instrument creation for the mobile and on the mobile, using a simple graphical interface and the notion of flow. Unfortunately, these projects are not updated as fast as the companies' mobiles and development kits. In the commercial field, RjDj is a sound application for the iPhone that allows users to access “scenes” -PureData patches- that transform sound that comes from the microphone and further manipulate it via the accelerometer and the touch screen.

Again the scheme here is changing fast and the various tools may be useful or not depending on the needs of the design. A full native development provides more flexibility but is more time consuming and complex.

Useful libraries at the moment are libpd[10], iPd[11], BuildingPdForiPhone[12].

---

[10] http://gitorious.org/pdlib

[11] http://code.google.com/p/ipd/

[12] http://puredata.info/docs/developer/BuildingPdForiPhone

3. Tangible tabletops and smartphones, shared control and collaboration

Interactive tangible tabletops are now drawing the attention of the digital musical instruments community. They seem to constitute a good platform for live music creation and experimentation. Their collaborative nature proves to be quite intriguing for the musical performance. The reactable was designed since the very beginning as a collaborative multi-user instrument and until this moment offers a quite intuitive approach to live digital music performance and improvisation. Other musical tabletops have also emerged, such as the Jam-O-Drum, The xenakis table or ANTracks.

Apart from music creation, tangible tabletops are definitely gaining territory as tools that promote collaboration and teamwork. They find uses in diverse fields, from collective business projects, museums and art installations to transportation control and medicine. On the other hand, mobiles are already researched as potentially interesting musical controllers with quite good results already, as seen in the previous chapter.

Although there has been an important effort, both in academic and commercial fields, to integrate smart-phones with tabletops for general purpose applications, musical integration has not been thoroughly investigated yet. ANTracks 2.0 [17] presents a musical instrument operating jointly on a multitouch tabletop and mobile phones. The interface is the same in both devices and the metaphor of ants moving in a grid is used to interact with pitch.

Figure 3.1: ANTracks 2.0 on a mobile phone and a tabletop

The interest of the present study is in using the affordances of the mobile as **an alternative way of controlling** the reactable. By just using the multiple touches and the accelerations as an input for the table, there are already **many degrees of continuous control** available. The existing objects do not have so many inputs. Neither they are **active** nor capable of providing some sort of feedback!

a) General purpose integration

Most of general purpose solutions deal with collaborative issues, but imply a sequential control paradigm as well. File sharing and distribution of the work in collaborative projects are prevailing. As for continuous control, the most frequent application seems to be the use of phones as controllers for larger displays, tabletops and games. Remote desktop control is possible by using the mobile as a mouse, while another approach is having the desktop entirely displayed on the mobile screen.

Gestures recognized from the accelerometer data have been researched as a way of integration of mobile phones with larger displays. MobiToss [18] implements a gesture based interface for 'throwing' videos on a large public screen and controlling the playback and effects. User-defined gestures are investigated in [19] under different settings (phone to phone, phone to tabletop, phone to public display).

Tactile feedback for text entry on a tabletop using mobiles is investigated in [20], concluding that feedback increases speed of typing, although mobile phones' soft keyboards are not a reliable solution.

PokerSurface [21] is a digital card game played with both mobile phones and a tabletop. It provides an interesting solution in the sense of extending the tabletop's capabilities. The mobile phones are used to resemble the cards in a simulation of a traditional poker game. Additionally, the accelerometers are used to hide or make visible the cards depending on the orientation of the phone. This gesture is much closer to the typical way of playing the game and it results much more natural according to the user tests. The privacy issue that is fundamental for this type of game is naturally solved, something impossible with just the tabletop.

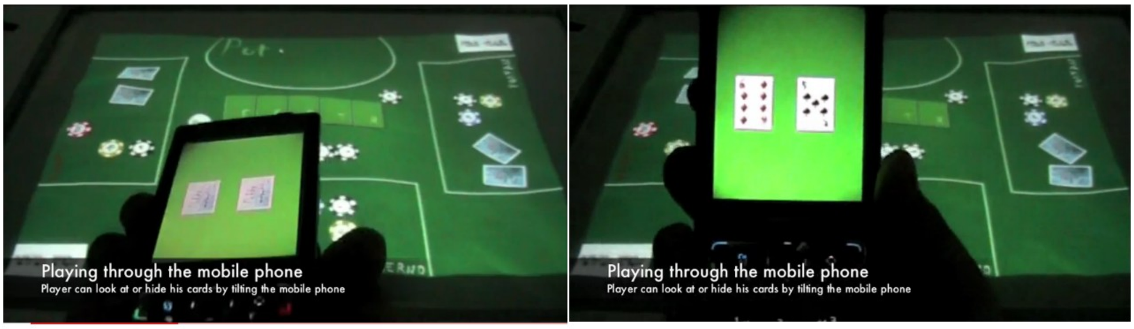


Figure 3.2: Use of accelerometers to hide/show the cards in PokerSurface

b) Communication issues

Apart from the conventional phone usage, the available options for steady communication between devices within a short range are the use of a Wireless Local Area Network, a Bluetooth connection and a recently proposed lower bandwidth optical connection! The camera and the screen of the tabletop along with the camera (predefined RGB patterns) and the flash of the mobile are used in [22] to easily establish an entirely optical connection, with less bandwidth than a WiFi or Bluetooth connection but totally secure. Bluetooth provides sufficient bandwidth but is mostly used for short ranges up to 10 meters. It also suffers from interference problems limiting the number of possible simultaneous users. WLAN also provides sufficient data rate and affords much greater distances, while there are more applications that facilitate the communication too. Therefore, WLAN seems suitable for the purposes of this study.

In particular, the protocol mostly used in real time communication for music is OSC. Open Sound Control [23] is an open message-based protocol developed for communication among computers, sound synthesizers, and other multimedia devices. It is based on a packet delivery model and is designed to be transport-independent, able to deliver over different media and telecommunication channels. This way, it solves the problems of MIDI addressing model and numeric precision limitations. In conjunction with UDP (User Datagram Protocol), a simple transmission protocol that drops packages instead of waiting and avoids error checking and correction, it seems a good solution for real time music.

Some applications that facilitate the communication worth mentioning are TUIOpad, a free app that sends multi-touch events using TUIO, the same protocol used in the reactable to treat messages from objects, as well as touchOSC, OSCemote and SonicLife. OSCulator is an environment running on Mac OS X that facilitates the connection with mobiles, as well as Wiimotes, tablets, mouse and keyboards. All of the above use OSC, most of them over UDP.

Another communicative issue is the fast and effortless establishment of a connection between devices. A method based on the shadow of the mobile on the table's surface together with the strength of a Bluetooth signal is presented in [24]. PhoneTouch [25] is a technique enabling file exchange as well as affection of the control of the shared device from the mobile one. Finger versus phone discrimination is achieved optically on the tabletop by instructing users to touch the shared surface with the edge of the phone and calculating the shape difference from a finger. Concurrently, accelerometer data is calculated on the phone for each touch of the device on the table's surface. This makes possible the recognition of each device's ID and consequently the ability to use multiple phones. However, the technique aims more in facilitating file exchange through a sequential type of control, thus it is not so interesting for the development of a real-time musical application, apart from the ability to recognize each mobile's ID.

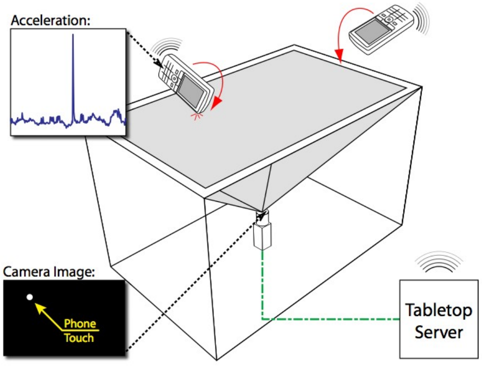


Figure 3.3: PhoneTouch functionality

A simpler solution, although not a sophisticated one, in the case of the reactable is to use the reacTIVision that tracks markers placed on the surface of the objects together with a new sticker on the back of the phone. Then a mobile with the sticker-identifier placed on the surface of the reactable could be identified as a mobile phone and establish the necessary connection.

c) Shared control and remote collaboration

Most traditional instruments are designed to be played by exactly one person, even though their ultimate goal is the musical collaboration with others. However, this collaboration is carried out in a `one player – one instrument´ or `one player – multiple instruments´ context most of the times, but we rarely see a `many players – one instrument´ situation. The traditional instrumentalist thus dives into collaboration by maintaining certain characteristics throughout the performance. These characteristics define in a way his personal identity, his ´voice´ that is related to some extent to timbral properties.

Delegation and the multi-threaded processes that many digital instruments take advantage of [27], exploit and use as a core feature are not present in acoustical ones. In pure musical terms, the distinction is fuzzy as different musical processes are some times sought for when still playing with just one traditional instrument, that is playing two or three 'voices' at the same time. This is at least perceived in a way as a multi-thread process, however the resulting timbre is always bound to the physical qualities of the instrument, something completely absent in digital instruments.

In terms of control however, traditional instrumentalists have clearly defined interfaces and boundaries (their actual instruments!) and who does what is normally not debatable. Digital instruments on the other hand can change their behaviour completely or automate/shift the control entirely with a single action. This is a huge challenge, as observed by the increasing amount of digital multi-user instruments appearing, with the reactable being a wonderful example. Sharing of control and collaboration becomes a core issue as spatial boundaries, as well as collaborative issues like leadership, are not anymore physically bounded, but rather in dynamic evolution. Roles change quickly and setting the boundaries becomes a constant task of collaboration and negotiation.

Since its very beginning, the reactable has been planned as a collaborative instrument. More complex collaborative models than the standard playing mode have been researched since the first stages of development. Various scenarios for both local and remote collaboration and participation are discussed in [26]. Two styles are mentioned regarding local collaboration, one where users define their territory by negotiating different regions of the table in a mode similar to a band. The second and more rich full in terms of collaboration and interaction, is when players share the whole space and play together, which however has to be planned and performed carefully in order to maintain its constructive nature as mentioned. Remote collaboration on the other hand was examined by a concert configuration of two reactables, one in Barcelona, Spain and the other in Linz, Austria playing together. Objects controlled by the distant table appeared on the surface as virtual objects and a visual projection informed the audience about the distant performers. An alternative participation scenario that is closer to the purposes of this thesis is mentioned, where members of the audience use a PDA equipped with a simple version of a reactable simulator and take part in the performance. Since then however, the evolution of the instrument was not really based on these collaborative scenarios.

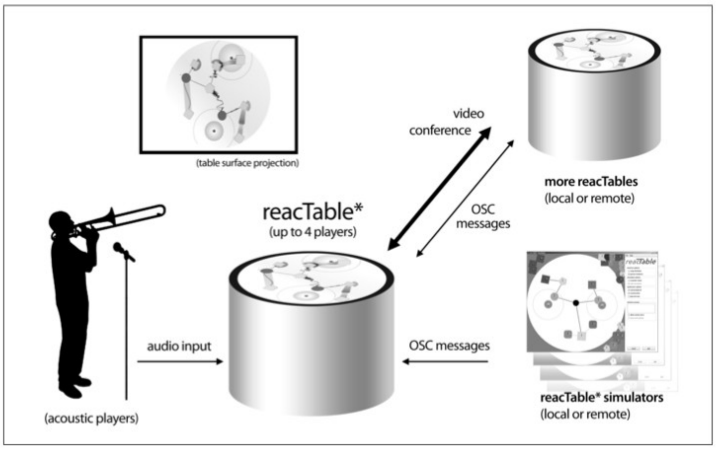


Figure 3.4: Local and remote collaboration scenarios for the reactable

This thesis is conceptually found in between these local and remote collaborative situations. The performer using the mobile is within a close range of the actual instrument and has the ability to move around. The network structure and tools used are for the purposes of a close distance communication, actually bounded by the WLAN range, but easily extendable for larger distances. The interest however is not in remote collaboration, but rather in this proximal interaction and the collaborative issues it implies. Roles in terms of control are to be distinct but musical ones are to be negotiated.

d) Expressiveness in realtime musical performance with digital instruments

Expressiveness can be a difficult or even impossible notion for a researcher to entirely grasp and define, while at the same time extremely intuitive and given for a musician or artist in general. J. S. Bach is reported to have said "*There's nothing remarkable about it. All one has to do is hit the right keys at the right time and the instrument plays itself.*"

In terms of digital instruments though, the expressive capabilities a performer has with his setup are even questioned or at least not taken as granted many times. In contrast to Bach's aforementioned quote, a spectator of a live laptop performance seems sometimes not to fully understand how the performer's actions affect the sound reaching his ears. A straightforward relation between the gestural control of the digital instrument and the produced sound is many times not even existing, or at least it cannot be easily perceived or learned. This can be partly attributed to the fact that digital instruments have as a real basic and easy to implement feature the delegation of actions, processes, behaviours. Many musical processes, simple or complex, are delegated and then reproduced automatically, without the need of constantly providing energy to the system. This can be a great deal of freedom for the musician but can be a contrasting effect for the spectator's perception as well. On the other hand digital controllers capable of complex, parallel and continuous control are not quite widespread, so many commercial products are based on the typical sequential control paradigm. This is somehow changing now with the increasing amount of mobile apps developed, with some of them focusing exactly on these new control inputs, such as the multiple touch control.

For the purposes of discussing interactive gestural instruments and with a specific focus to mobile devices [6], Atau Tanaka regards expressiveness as "*specific musical affordances of an instrument that allow the musician performing on the instrument to artfully and reliably articulate sound output of varying nature that communicates musical intent, energy, and emotion to the listener. The success of expressivity resides not just in the effectiveness of communication, but in the sense of agency that the system gives back to the performer.*"

In [30] expressiveness is said to be "*the capacity to convey an emotion, a sentiment, a message, and many other things. It can take place at various levels, from the macroscopic to the microscopic scale. In the case of musical performance, expressiveness can be associated with physical gestures, choreographic aspects or the sounds resulting from physical gestures [..]* ***tends to involve using specific gestures to obtain an expressive sound rather than performing expressive gestures***".

In acoustical instruments gestures result from compromises between the acoustics of the instrument and the physiology of the human body. The traditional luthier does make a lot of choices, but is often closely bounded by the materials he is working with. He normally concentrates more on delicate and detailed issues of the instrument development than he is asked to make fundamental decisions about how the players' actions are going to be translated into sound. These issues seem to be already solved in traditional instruments through a constant evolution of hundreds or thousands of years.

In digital instruments though, the mapping between sound and gesture comes down to design choices. Physical constraints do exist in the complex control-mapping-sound production chain. The character of the instrument however (for example its basic functionality, how/if the sound is going to be synthesized, the resulting timbre) can be defined more by the digital luthier. An extremely abstract or even non-meaningful, boring or tiring mapping is not only possible but in fact really easy to make and maybe quite often to see.

In the work of Arfib, Couturier and Kessous, expressiveness is related to identity and adaptability. The identity of an instrument lies at the note and phrasing level, while its adaptability is the potential of blending into or following other musical styles and pallets, as well as the ability to 'emerge'. Under a similar scope in [27, Chapter 7.14], expressiveness in realtime performance is seen as deviation from the predefined context, as well as musical diversity, the ability of an instrument to allow the performer to follow diverse musical paths. This diversity control is thoroughly examined in the same work, with an interesting discussion on the relative benefits of digital instruments. A quite intuitive but sometimes easy to forget conclusion is also highlighted. Musical instruments and machines in general, should be judged on their ability to transmit human expressiveness, they have nothing to express however, thus they cannot be expressive.

Visual considerations are present in both aforementioned studies, as visual feedback is a clear advantage of digital instruments. However, how expressive one can be with an instrument does not usually depend on the visual feedback, or at least it is not the most important factor. It can account nevertheless for the comprehensibility of the design. It can help the people in the crowd to immerse themselves in an audiovisual experience (mobiles can also play a role here). It can help retain their attention, specially if they can make a connection between the visuals and the sound. The visual feedback can consist of the illustration of low-level parameters, interpreted mid-level concepts or metaphors, some actual representation of the sound generated or even a completely abstract representation. Whatever it is, it has to fulfil the needs of the specific design and to help the player engage in the experience with the instrument. Synchronization issues are quite important here.

Another common point in the two studies is a control example given as interesting, intuitive and somehow expressive. That is the control of the depended parameters resonant frequency and bandwidth of a filter with the highly-coupled parameters of a joystick .

4. One voice singing in between the reactable and a mobile

The word telephone (τηλέφωνο) comes from the Greek words "τηλέ", that means distant, coming from a distance and "φωνή", that is voice. A literal translation could be distant voice. Although this kind of metaphor was not explicitly sought for at any time during the process of the design, it seems that it is quite suitable for the work carried out here. The resulted system is a voice synthesizer based on formant synthesis controlled by the mobile device and some reactable objects simultaneously. So, it can be said that it is somehow a 'distant voice', something that could be actually considered literal instead of a metaphor. However, this is a clue that came afterwards, as the whole design process was not based on such an idea but was aiming instead for an intuitive and expressive proposal for integrating with the reactable, baring in mind the affordances of the mobile platforms as examined in the previous chapter.

Every procedure of crafting a musical instrument or interface, apart from clearly engineering and technological aspects relates also to psychological, philosophical, conceptual, musicological and musical aspects [27, Chapter 1.3]. As stated in [28] "*Music interface construction proceeds as more art than science, and possibly this is the only way it can be done*". However, the technological tools are our main means of work and given the incompatibility scheme we have already mentioned, the design pretty much depends on the affordances of the devices and the development tools available. Nevertheless, the design concept and goal has to be always on top of the technological restrictions, if it is to be rigid and able to survive and evolve under the given conditions.

The proposed system is a voice synthesizer that aims not to be extremely realistic, but to exploit instead the potential of mobiles for providing a meaningful and intuitive integrated app between the two platforms. The proposed design is to be judged in terms of the sense of expressiveness it can provide to the player, as well as its ease of approach and intuitiveness. The goal is to have an intuitive and intriguing tool to be able to excite the performance in a simple way, given the small surface of a mobile and moreover its nature, that is not ideal for spending a lot of time. The sharing of control between the mobile and the tabletop is a core issue and is examined under different conditions and mappings. Generally, the tabletop is responsible for the sound synthesis while the mobile is a controller. However, the tabletop also controls part of the synthesis and the mobile can also be more than a mere controller by providing for example some information about the running state of the reactable. The implementation consists of a PureData patch running on the server side that is responsible for the whole sound synthesis and receives raw control data in the form of OSC messages. On the mobile phone side, an interface is developed in iOS using the iPhoneSDK and the VVOSC framework[13] to handle the communication.

a) Formant synthesis and the Klatt model

Having as main goal to explore and exploit the sensory features of mobiles that are not available in the standard reactable installation, the selection of the digital sound synthesis technique seems to be an important factor. What is needed is a sound generator with many possible inputs to experiment with and a meaningful outcome for these inputs. Curtis Roads

---

[13] http://code.google.com/p/vvopensource/

[15] provides a good review and some sort of classification of the various digital sound synthesis methods. Formant synthesis, a method belonging in the general category of physical modelling is chosen as a first interesting step to explore the integration possibilities. Physical modelling is somehow suitable when dealing with gestures, while formant synthesis results intuitive for a user to control as stated in C.Roads' “the computer music tutorial” :

> *¨Formant synthesis gives musicians a direct handle on one of the most important sound signatures [..] the spacing and amplitude of spectral peaks.¨*

With this type of synthesis all vowels can be generated by just controlling the amplitude of voicing, the fundamental frequency F0 and the first three formants and their amplitudes. To simulate a singing voice in an electronic music context, vowels and vowel-like sounds seem to be enough.

The idea is to share the control of a voice, maybe a soloing one, between a performer on the reactable and one standing away controlling a mobile. The cascade/parallel formant synthesizer by D.Klatt [16] is taken as a basis. Although the cascade format is closer to a realistic simulation of the vocal tract and there is no need to adjust the amplitudes of the formant filters, the sound output is less predictable when navigating in a 2-D space with axes the first and the second formant. The parallel configuration can also produce fricative and plosive consonants. Any possible simplification is made while maintaining a satisfying vowel-like outcome, for the sake of simplicity and to focus on the control aspects.

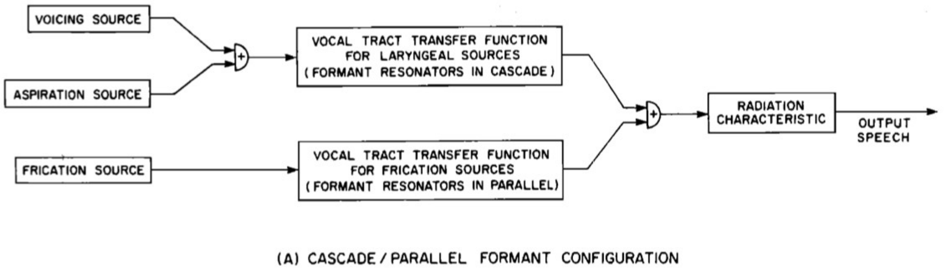


Figure 4.1: Generic diagramm of the Klatt formant voice synthesizer

Vowel templates are normally available for a specific language and for the basic vowels (A,O,U,E,I). The standard process is the measurement of the formant frequencies for a large amount of human speakers under various contexts and the extraction of means that form the templates. Depending on the measurement however, researchers tend to categorize speakers differently, with the most frequent case being the separation between male, female and child. A more 'musical' categorization seems more suitable for our purposes and the templates chosen are the ones from the Csound manual[14] that give the formant characteristics for different kinds of singers:

<u>Soprano</u>

Vowel	**a**					**e**					**i**				
Freq (Hz)	800	1150	2900	3900	4950	350	2000	2800	3600	4950	270	2140	2830	3900	4950
Amp (dB)	0	-6	-32	-20	-50	0	-20	-15	-40	-56	0	-12	-26	-26	-44
Bw (Hz)	80	90	120	130	140	60	100	120	150	200	60	90	100	120	120

---

[14] http://mitpress.mit.edu/e-books/csound/csoundmanual/Appendices/table3.html

o					u				
450	800	2830	3800	4950	325	700	2700	3800	4950
0	-11	-22	-22	-50	0	-16	-35	-40	-60
70	80	100	130	135	50	60	170	180	200

Bass

Vowel	a					e					i				
Freq (Hz)	600	1040	2250	2450	2750	400	1620	2400	2800	3100	250	1750	2600	3050	3340
Amp (dB)	0	-7	-9	-9	-20	0	-12	-9	-12	-18	0	-30	-16	-22	-28
Bw (Hz)	60	70	110	120	130	40	80	100	120	120	60	90	100	120	120

o					u				
400	750	2400	2600	2900	350	600	2400	2675	2950
0	-11	-21	-20	-40	0	-20	-32	-28	-36
40	80	100	120	120	40	80	100	120	120

Tenor

Vowel	a					e					i				
Freq (Hz)	650	1080	2650	2900	3250	400	1700	2600	3200	3580	290	1870	2800	3250	3540
Amp (dB)	0	-6	-7	-8	-22	0	-14	-12	-14	-20	0	-15	-18	-20	-30
Bw (Hz)	80	90	120	130	140	70	80	100	120	120	40	90	100	120	120

o					u				
400	800	2600	2800	3000	350	600	2700	2900	3300
0	-10	-12	-12	-26	0	-20	-17	-14	-26
40	80	100	120	120	40	60	100	120	120

Alto

Vowel	a					e					i				
Freq (Hz)	800	1150	2800	3350	4950	400	1600	2700	3300	4950	350	1700	2700	3700	4950
Amp (dB)	0	-4	-20	-36	-60	0	-24	-30	-35	-60	0	-20	-30	-36	-60
Bw (Hz)	80	90	120	130	140	60	80	120	150	200	50	100	120	150	200

o					u				
450	800	2830	3500	4950	325	700	2530	3500	4950
0	-9	-16	-28	-55	0	-12	-30	-40	-64
70	80	100	130	135	50	60	170	180	200

Contratenor

Vowel	a					e					i				
Freq (Hz)	660	1120	2750	3000	3350	440	1800	2700	3000	3300	270	1850	2900	3350	3590
Amp (dB)	0	-6	-23	-24	-38	0	-14	-18	-20	-20	0	-24	-24	-36	-36
Bw (Hz)	80	90	120	130	140	70	80	100	120	120	40	90	100	120	120

o					u				
430	820	2700	3000	3300	370	630	2750	3000	3400
0	-10	-26	-22	-34	0	-20	-23	-30	-34
40	80	100	120	120	40	60	100	120	120

b) Developing the synthesizer and mapping the parameters

The reactable controls the fundamental frequency and the gain. These reactable controls exist in many objects already (see Fig. 4.1) and seem to be perceived quite intuitively, as the pitch we hear seems to have an elliptical mental representation simulated by this rotary behaviour. The amplitude is also controlled easily, by a finger lowering a semicircular slider for lower gains. On the other hand, timbre navigation is a bit of open issue for the instrument and is the most important 'handle' that formant synthesis provides, a synthesis technique not yet implemented on the instrument. Therefore, the mobile is made responsible for controlling the voice timbre. It does so by being inspired by this interesting example mentioned in the previous chapter about controlling a filter's resonant frequency and bandwidth with a joystick. Here actually, the touch screen is considered a vowel space where you navigate with a finger, so the finger position affects multiple filters and bandwidths at the same time. This is the most frequent implementation of realtime formant synthesizers, none of which was commercialized though. Apart from the timbre navigation, the vibrato characteristics are also being controlled by the mobile device. Other features such as breathiness and nasalization have been tried out as well.

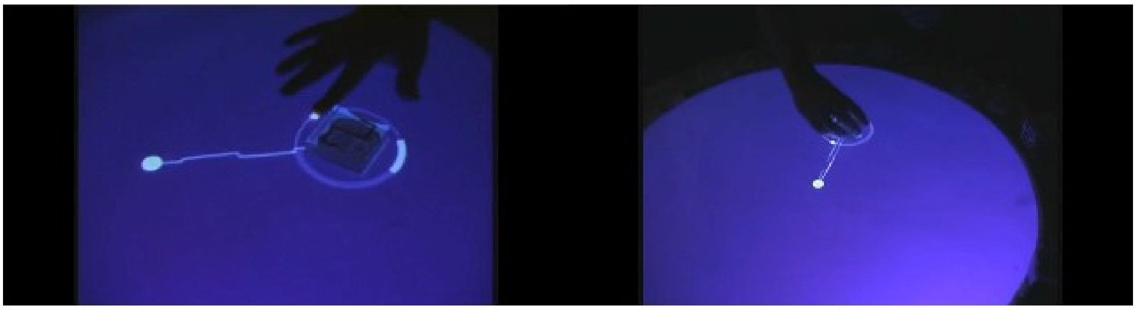


Figure 4.2: Controlling a reactable object by touching with a finger (left) and by rotating (right)

The synthesizer is based on three formant filters in parallel configuration. A phasor source simulates the voicing while an addition of a variable amount of pink noise simulates the noise source, that is needed for aspirated sounds and to add a kind of breathiness to the pure vowel. The first step one has to take is to choose a singer with the sub-patch seen in figure 4.3.

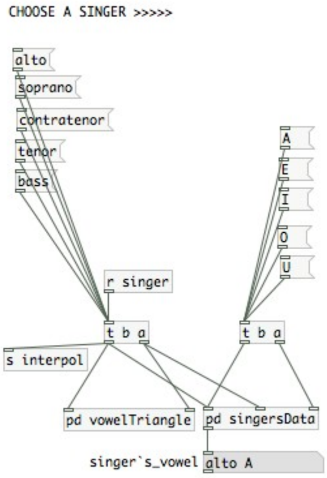


Figure 4.3: Selection of a vowel space / singer

The choice of singer defines the vowel space where the performer is going to navigate with his first finger that touches the screen. This is the basic form of control in general. From the templates given in the previous chapter, the vowel space limits are set for the navigation, as well as the reference values for the interpolation process. This process is needed to extract further formant frequencies and the bandwidths and amplitudes, apart from the first and second formant frequency that are controlled by the (x,y) position of the first finger on the screen (Fig.4.4).

Formant resonators used are two-pole filters and the incoming control values affect their resonant frequencies, bandwidths and amplitudes. The two first formants are enough to produce a comprehensible vowel, although the sound quality is not so satisfying. A cascade synthesizer with just two controls of formant frequency and steady bandwidths mapped directly to the x-y coordinates of the 2D navigation space is already quite fun to play with. For more convincing or realistic results though, further values have to be interpolated from the current cursor state, mainly the third formant frequency and bandwidth, as well as the gains of all formants.

Figure 4.4: Simulation of the basic interface for the mobile application: the vowel space of navigation.

The process that follows is a variant of linear interpolation based on the vertices of a triangle. The 2D space defined by the first and second formant frequencies in which we are about to navigate, is pretty well defined for each singer/talker from a triangle formed by the exterior vowels \`a\`,\`i\`,\`u\`. This triangle is a common tool for philologists, who denominate the axis corresponding to F1 as vowel aperture and the axis of F2 as articulation point. The third formant is said to be mainly affected by the height of the tongue tip. This triangle is regarded as the reference of some vowel space and when calculated versus a third value defines a unique plane, the plane on which we interpolate.

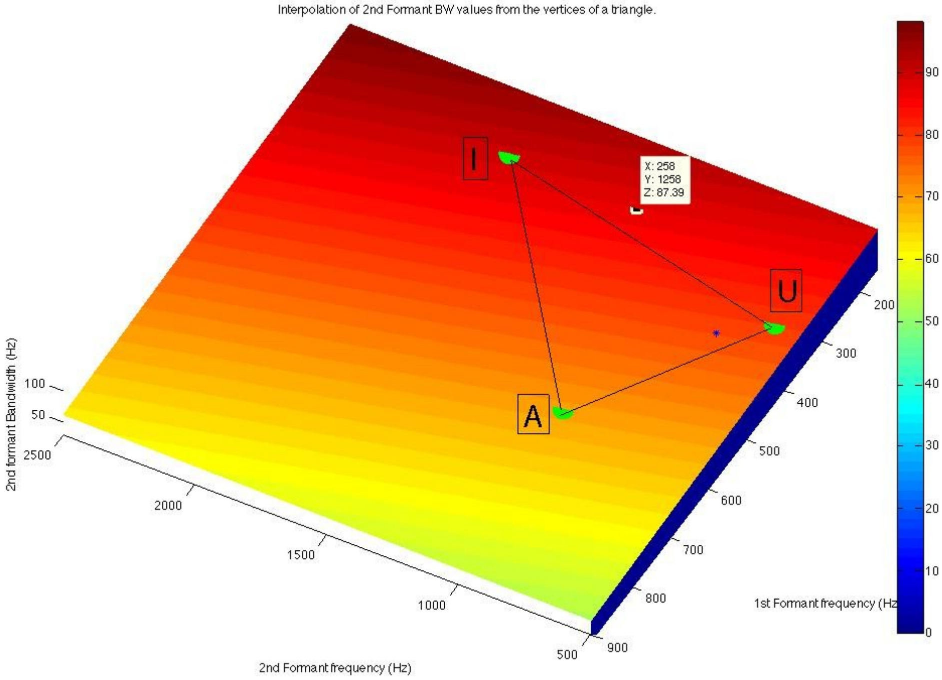


Figure 4.5: Matlab demonstration of the interpolation function for a bass singer

So the interpolation plane of the second formant bandwidth for example is defined by:

$$BW_2(U) = A_x F_1(U) + B_x F_2(U) + C_x$$
$$BW_2(A) = A_x F_1(A) + B_x F_2(A) + C_x$$
$$BW_2(I) = A_x F_1(I) + B_x F_2(I) + C_x$$

where F_1, F_2 are the first and second formant frequencies,

BW_2 the second formant bandwidth.

After solving for A_x, B_x, C_x , the bandwidth of the second formant is simply calculated as:

$$BW_2(P) = A_x F_1(P) + B_x F_2(P) + C_x$$

for any point P with coordinates F_1 and F_2.

The entire process is implemented in PureData. In figure 4.5 it is demonstrated in MATLAB for the case of a bass singer and the bandwidth values of the second formant. Comparing the values of the plot with the ones given in the templates for the two basic vowels left \`e\`,\`o\`, some difference of 2-3Hz is found for \`e\`, while no difference is observed for \`o\`. The accuracy is thought to be more than satisfying.

With the use of three formant frequencies, gains and bandwidths the sound outcome is already pretty vowel-like. Furthermore, research in speech recognition points out that vowels are mostly recognized by these three formants, as the higher ones are more obscure or even absent in many cases. However, the sound output is not equally satisfying throughout all frequency ranges that a real singer can handle. Singing involves complex processes of adaptation to the context, resulting in a delicate control of formant regions and bandwidths depending on the musical context and the purposes of the performance. Singer´s formant is a well known and studied technique of mostly male singers that when soloing they 'create' a formant around 3kHz that does not exist in speech. This is achieved by the merging of neighbouring formants.

The basic functionality of the formant synthesizer can be seen in figure 4.6. Blue sliders denote the parameters controlled by the reactable while the orange ones the parameters controlled by the mobile device. As it is, the patch can be controlled by a custom mobile interface, as well as the applications TUIOpad[15] and touchOSC[16]. The connection with the tangible tabletop is achieved by a UDP connection over Wi-Fi. After a long iterative process of development, try-out and re-design, the basic control parameters of the synthesizer were carefully mapped to the inputs of the mobile and the reactable. The basic control is the navigation inside the selected vowel space and it is achieved in two different ways. Either with the first finger that touches the screen of an interface in a single control area, or with whichever finger that touches a predefined area that is the main view but not covers the whole screen. The difference is actually in the way the controls of the second finger are affected. If there is a separate area for it, like a small vertical slider that does not need a lot of space, the design seems more comprehensible. However, by just following the order in which the fingers touch a single-area interface to distinguish one finger from the other, a sense of the controls being coupled is created. They are coupled actually, as for the second finger to get to play, the first finger has to be in touch with the screen.

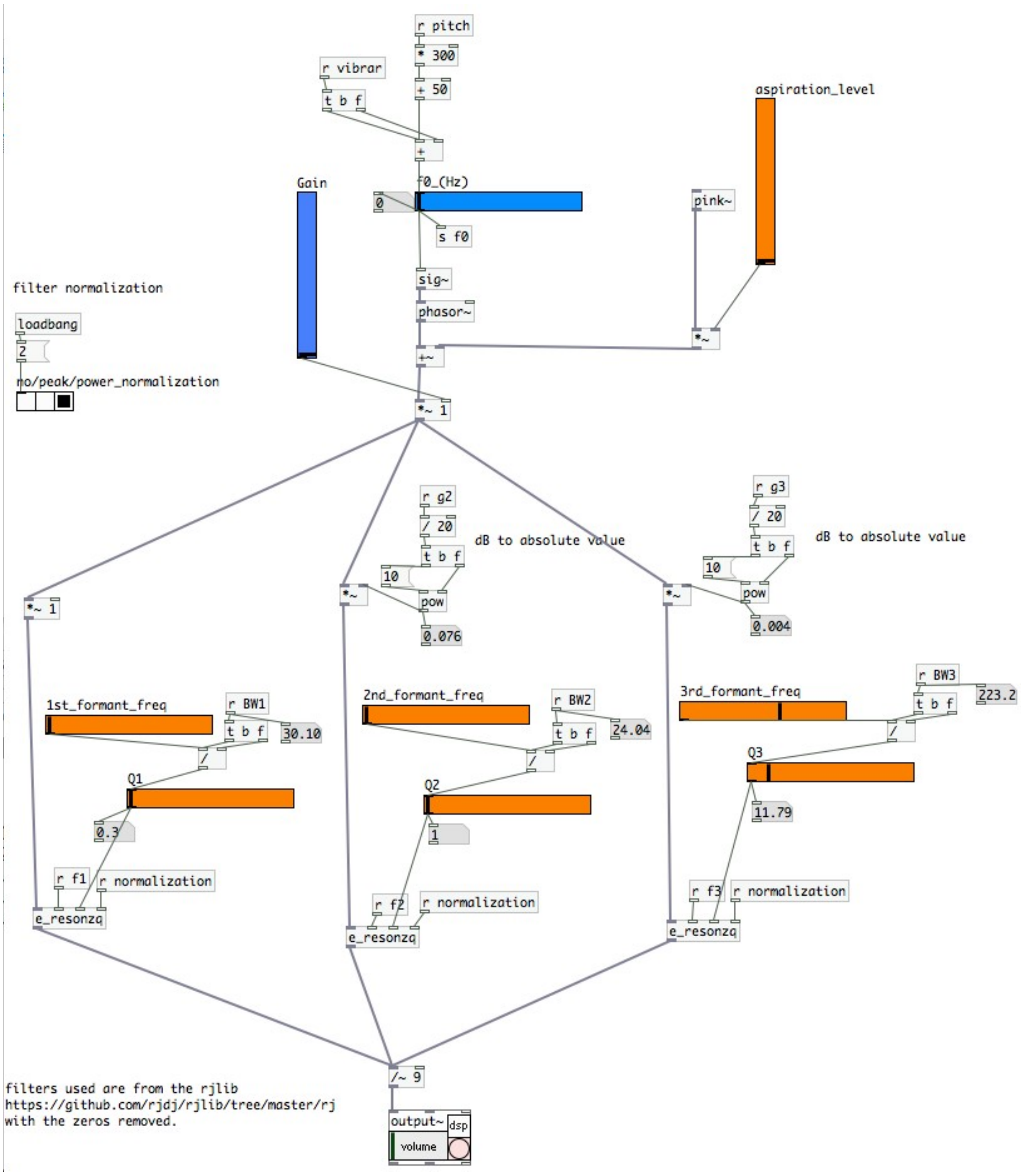
Furthermore, a simple toggle is used to change the way the second finger that touches the screen affects the sound generation. One mode simply re-triggers the process in the same way, resulting in more rhythmic patterns. In the second state the second finger controls vibrato and the third finger the amount of noise present in the source. The gain was also assigned to this control with quite good results. With the use of a single vertical slider, the performer can control his dynamics, something quite easy to understand and control.

---

[15] http://code.google.com/p/tuiopad/

[16] http://hexler.net/software/touchosc

An important feature is the control of time events by the touches. Although a simple ON-OFF functionality using a toggle gives good results, the experience gets much more interesting when the triggering of any sound event is directly mapped to the touches, meaning that for the instrument to sound a finger has to be touching the screen. This gives in some way the sense that energy is needed for the instrument to come into action and is helpful in engaging the player. Vibrato is simulated as a simple modulation process with two input parameters, the vibrato frequency and amount. Out of the various mappings that were tried out for the vibrato, the ones that felt better playing were the use of accelerometers as well as the use of the second finger in the case where finger touches are coupled (there is the need for a first finger also). The accelerometers result quite interesting when the two vibrato parameters are controlled by the rotation of the mobile in two different directions. The vibrato results somehow more difficult and complicated to control, but at the same time easy to discover and intriguing. A breathiness/aspiration effect is also achieved by simply setting the amount of pink noise added to the source and was controlled by a simple vertical slider.



filters used are from the rjlib
https://github.com/rjdj/rjlib/tree/master/rj
with the zeros removed.

Figure 4.6: Basic functionality of the formant synthesizer developed in PureData

c) Interface considerations

The tabletop interface was developed using the Musical Tabletop Coding Framework[17] developed by Carles Julià. It consists of one or two simple objects per mobile. In one case an object is used for the fundamental frequency of the voice and another one for its gain, with all the control carried out by the rotary behaviour. In the other case, a single object is used per mobile and the fundamental frequency is still controlled by rotating while the gain is assigned to the slider next to the object and it is controlled by a finger. The whole interface is developed for the sake of simplicity and the purposes of this prototype, as a careful connection between the formant synthesizer and all the existing reactable objects is time consuming and needs to be done carefully.

The navigation space of the mobile application is already set as the 2D plane defined by the first and second formant frequency. It seems to be quite intuitive to navigate from one vowel to another, even when you have no visual feedback of the current vowel position. The first view that appears on the mobile user's screen is a simple view for the choice of the singer's characteristics. The options are bass, soprano, alto, tenor and contratenor, but with the use of different templates other options can be added, such as male, female or child. After the user chooses a singer, the main playing view appears, consisting of a simple control screen with the basic vowels placed across the two-dimensional plane. It seems to be extremely easy and immediate for a user to grasp the connection between the finger position and the resulting vowel sound with this kind of interface.

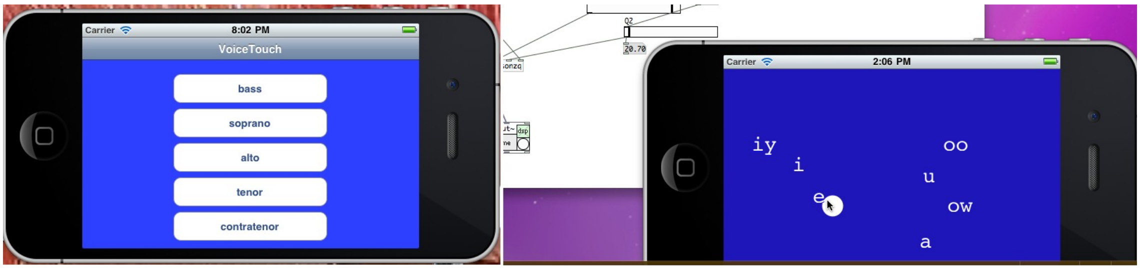


Figure 4.7: the view for setting the singer´s vowel space (left) and the main playing view (right)

The interest does not seem to lie on how realistic a certain vowel comes out, but rather on how you can simply move along this space with a finger gesture. Even small phrases full of vowels are possible with a bit of practice. Phrases like 'how are you' or ' I love you' can even be intelligible without actually having heard the consonants in between. It just seems that this finger gesture is somehow a good approximation of the formant evolution, or at least it is random enough -or human enough?- to provide a realistic sense of transition from one vowel to another.

However, when trying to adjust the vowel positions on the display, the pitch-formant dependency results in the vowels shifting, when the 'voice' does not sing in the central frequency region of the chosen singer. The display however is quite small and the problem can be somehow bypassed by adjusting the frequency regions and the letter display or animation.

In contrary, if the whole chain driven by one performer with a mobile and another one on the reactable is considered as a single instrument, the mobile could be seen as a controller while the reactable would be the main body of the instrument. That would evoke this particularly delicate but extremely interesting subject of exciter-resonator coupling, present in most acoustical instruments, absent in many digital ones and certainly a great handle to engage

---

[17] http://mtg.upf.edu/research/interaction/technologies?p=Musical%20Tabletop%20Coding %20Framework

players for many successful instruments. In digital sound synthesis this can be observed in physical modelling techniques as stated by Curtis Roads in his computer music tutorial [15]:

> *¨When the exciter and the resonator are coupled (the vibration of the resonating part feeds back to the exciting part), a variety and subtlety of sound is created. By modelling this, PhM techniques provide a sense of gesture, missing in abstract synthesis methods.¨*

Under the former assumption of the mobile being a controller and the reactable the 'resonator', an important visual clue could be given to the performer. The change in pitch from the reactable could be translated in an adaptation of the mobile user´s interface. The hypothesis is that this could create a certain feeling of 'connectedness' with the tabletop performer. Another solution is just to adjust the display in a way that letters look bigger and thus cover larger areas. With a fine adjustment combined with carefully setting the frequency limits of each singer -maybe restricting the regions much more than those of a real singer- there would be no problem of 'shifting' vowels.

5. Experimenting and concluding

While there is not any obvious parameter for direct evaluation, there are still many open issues that lead to interesting experimentation possibilities in the fields of musical collaboration and shared control. While not trying to discard usability evaluation, a more informal process that was going to provide active feedback was sought. Usability evaluation results very helpful in various stages of the development of an interface, it can easily mislead however if done without extensive thought. As stated in [31]:

> *¨The choice of the evaluation methodology -if any- must arise from and be appropriate for the actual problem or research question under consideration.¨*

User tests were discarded due to the former, as well as the time shortage. Two different procedures took place instead. The first one was the trial of the application using a single mobile device and the reactable by two people, myself as the designer and 'experienced' user of the instrument and a subject with no prior experience of playing with the mobile. The other procedure consisted of three performers with mobile phones playing simultaneously with one performer on the reactable. Each mobile and its corresponding reactable objects represented a voice, in a way that the reactable performer acted as a maestro in our improvised electronic choir.

a) One mobile phone and the reactable

Different settings and mappings were tried out during this process. First, a single player controlled the whole system with the two hands, one on the reactable object controlling the frequency and the other one on the mobile controlling the timbre. Although a bit out of scope, the system seemed suitable for bimanual experimentation. Then, the ordinary setup was established, with one performer on the mobile and another one on the reactable. The timbre navigation resulted easy enough for the unexperienced performer to grasp in a few minutes. Then the various mapping schemes mentioned in the previous chapter were tried out. The most intriguing one seemed to be the control of vibrato with the use of accelerometer measurements in two axes, apart from the standard navigation across the vowel space with the finger.

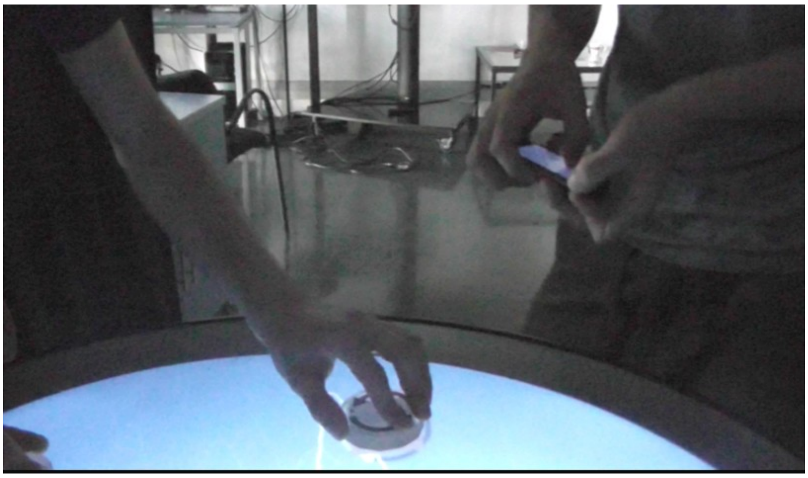


Figure 5.1: Shared control of the application between a performer with a mobile phone and another with the reactable.

A demo can be found in http://www.youtube.com/watch?v=l0Zp9-qQrwQ&.

Two different schemes were tested for the mapping of the vibrato parameters to the accelerometers of the mobile. The first one implied a direct dependence between the fundamental frequency and the mobile's rotation degree. It resulted in having to vibrate the mobile to achieve a vibrating voice and although quite nice visually, it was somehow difficult to control and produce a realistic vibrato. The second mode was to assign the vibrato depth and amount values to the mobile's rotation, such as a tilt in the mobile triggered a certain vibrato value. It resulted less realistic than the first case -a certain rotation angle meant certain vibrato characteristics, so there was no need for doing something to sustain this vibrato behaviour-, but it also seemed easier to control and explore.

b) “Our little choir”

Three mobile devices – two iPodTouch and an iPad – representing three electronic voices were used for this improvised electronic choir. Each mobile established a connection with one specific reactable object using a different port and then sent control data using a custom interface created with the touchOSC application. The fundamental frequency was always controlled by rotating this object as usual, while the gain of each mobile was assigned to both the mobile itself and the object's slider during the process. The reactable performer had thus an important and central role, to control the melodic evolution and the consonance of the voices. On the other hand, mobiles controlled the timbre and vibrato of their corresponding voice, as well as its breathiness and amplitude. Five subjects took part in this performance with the two of them having some previous experience playing with the instrument. The reactable player was the same throughout the performance while mobile players changed roles. The voices were distributed to different frequency ranges, with the bass singer assigned to the iPad and a tenor and soprano singer assigned to the two iPodTouch. The players were given simple directions about the application's functionality and were then left to improvise.

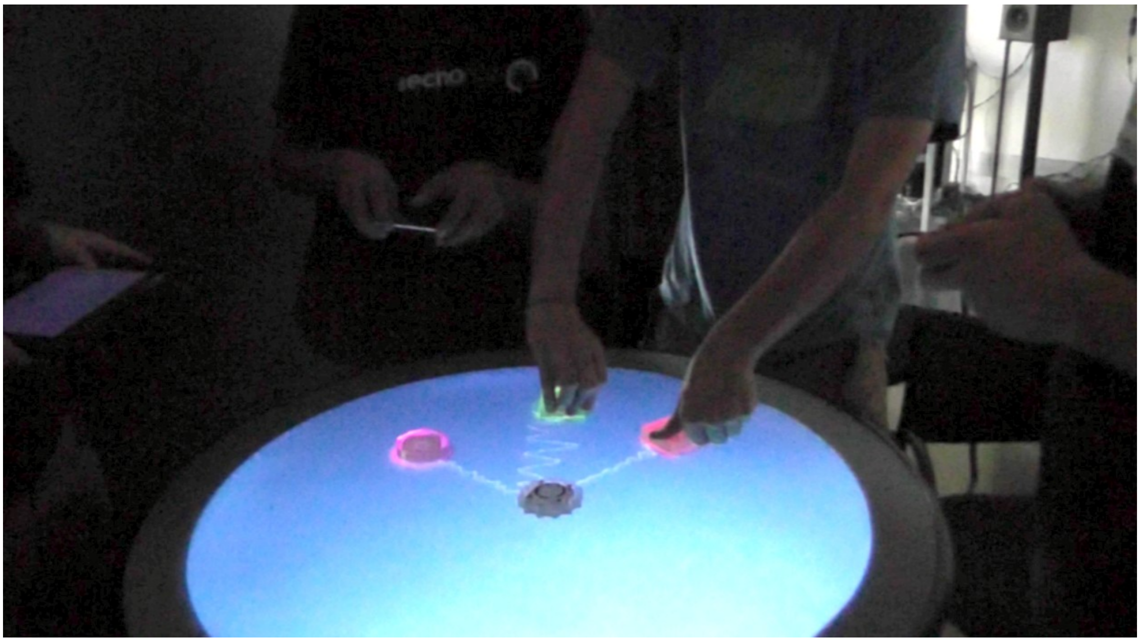


Figure 5.2: three tele-'phones' and the reactable as a maestro

A demo can be found in http://vimeo.com/groups/main/videos/25966968, while a whole song entirely improvised in http://www.youtube.com/watch?v=2Og-d5NcWTA. The feedback was quite encouraging as the players engaged easily in the experience and were able to play collectively after a short time. They seemed to enjoy the freedom of movement the mobiles provide. It is easier to move and dance while playing with a mobile than when playing with just the reactable. The vibrato seems to be quite substantial for the user to perceive the output as a voice. Furthermore, with the mapping of its parameters to the accelerometer values, that are somehow unstable -if a gyroscope is not present -, the voice becomes more natural as it always includes some amount of vibrato, even when the mobile is at an horizontal position. Some conclusions that came out of this process were the preference for the control of the gain of each voice by the mobile performer himself and not by the reactable, as well as some suggestions about the interface structure and the size of the vertical sliders controlling secondary parameters such as breathiness.

c) Conclusions and proposals for the future

The formant synthesizer was used here as a case study and showed that the integration of musical tangible tabletops and smartphones is quite promising and opens up a lot of possible directions for experimentation. The degrees of freedom that the mobiles provide for control can constitute musically interesting inputs for the reactable in a variety of ways and substantially increase its control capabilities. As for the formant synthesizer, the results are promising as well and a potential integration with the standard reactable installation could be useful. A standalone application is also part of the future plans.

For a generalized framework of integration of tangible tabletops with mobile phones, extensive work has to be carried out to account for suitable ports and means of connecting diverse applications under various contexts. The results of this study though show that such an integration framework could substantially upgrade the control capabilities of the reactable, as well as the whole user experience with it. Live performance with digital musical instruments is in constant evolution and the advanced technology of modern mobiles can offer a lot in the musical interaction. The extensive amount of mobile apps available can also form part of an interactive installation and the reactable can most certainly play a central role, as it did in “our little choir”. What is needed is the creation of suitable generic ports for the reactable to be able to communicate with mobiles under various musical contexts. A mobile can be seen as an external instrument and connect to the reactable with a simple wireless connection. Then, the external signal can be treated as usual on the surface of the reactable. Of course the diversity of the mobile apps is huge and not all of them are interesting, neither they can be treated equally. Nevertheless, the development of a generic framework that would facilitate the integration of the reactable with a large amount of mobile apps is possible, or at least the creation of a layer on top of which diverse interaction modes could function in conjunction with the reactable.

Bibliography

1. JORDÀ S., GEIGER G., ALONSO M., KALTENBRUNNER M. "The reacTable: Exploring the Synergy between Live Music Performance and Tabletop Tangible Interfaces", TEI'07, Baton Rouge, LA, USA.
2. BOSI, M. "Extending Physical Computing On The Reactable", Master Thesis (2009), Univeristat Pompeu Fabra ,Barcelona, Spain .
3. KALTENBRUNNER M., BENCINA R. "reacTIVision: A Computer-Vision Framework for Table- Based Tangible Interaction", TEI '07, Baton Rouge, Louisiana, USA.
4. BEHRENDT F., TANAKA A., GAYE L., HOLMQUIST L.E. "Mobile Music Technology: Report on an Emerging Community" , NIME'06, Paris, France.
5. TANAKA A. "Mobile Music Making", NIME'04, Hamamatsu, Japan.
6. TANAKA A. "Mapping Out Instruments, Affordances, and Mobiles", NIME '10, Sydney, Australia
7. http://www.reactable.com/products/mobile/ , last checked 4/4/2011
8. WANG G. "Designing Smule's iPhone Ocarina" , NIME'09, Pittsburgh, PA. USA.
9. WANG G., DAHL L. "Sound Bounce: Physical Metaphors in Designing Mobile Music Performance", NIME '10, Sydney, Australia.
10. ESSL G., ROHS M. "CaMus2 -- Optical Flow and Collaboration in Camera Phone Music Performance" , NIME07, New York, NY, USA.
11. KRATZ S., ESSL G., ROHS M. "Use the Force (or something) - Pressure and Pressure- Like Input for Mobile Music Performance" , NIME'10, Sydney, Australia.
12. YÜKSEL K.A., KETADBAR H., JAHNBEKAM A. "MagiMusic: Using Embedded Compass (Magnetic) Sensor for Touch-less Gesture Based Interaction with Digital Music Instruments in Mobile Devices", TEI '11, Funchal, Portugal.
13. OH J., WANG G., BRYAN N.J., HERRERA J. "MoMu: A Mobile Music Toolkit", NIME '10, Sydney, Australia.
14. MÜLLER A., ESSL G. "Designing Mobile Musical Instruments and Environments with urMus", NIME '10, Sydney, Australia.
15. ROADS C. "the computer music tutorial", Part II: Sound Synthesis, p.81-346. The MIT Press (1996), London, England.
16. KLATT D. "Software for a cascade/parallel formant synthesizer", Journal Of Acoustic Society of America (1980), p. 971-995.

17. RECKTER H., SCHULZ F., WÖLDECKE B., GEIGER C. “ANTracks 2.0 - Generative Music on Multiple Multitouch Devices”, NIME '10, Sydney, Australia.
18. COULTON P., SCHEIBLE J., OJALA T. “MobiToss: A Novel gesture based interface for creating and sharing mobile multimedia art on large public displays”, MM '08, Vancouver, Canada.
19. KRAY C., NESBITT D., DAWSON J., ROHS M. “User-Defined Gestures for Connecting Mobile Phones, Public Displays and Tabletops”, MobileHCI’10, Lisbon, Portugal.
20. BREWSTER S. MCADAM C. “Distal Tactile Feedback for Text Entry on Tabletop Computers”, HCI '09, Cambridge, UK.
21. PARVAHAN P., AHRENS B., SCHMIDT A., SHIRAZI A.S., DÖRING T. “Poker Surface: Combining a Multi-Touch Table and Mobile Phones in Interactive Card Games”, MobileHCI'09, Bonn, Germany.
22. BOLL S., HESSELMANN T., HENZE N. “FlashLight: Optical Communication between Mobile Phones and Interactive Tabletops”, ITS'10, Saarbrucken, Germany.
23. WRIGHT M. “Open SoundControl: A New Protocol for Communicating with Sound Synthesizers”, Proceedings of the 1997 International Computer Music Conference.
24. KLINKER G., ECHTLER F. “Tracking Mobile Phones on Interactive Tabletops”, MEIS'08, München, Germany.
25. RUKZIO E., GELLERSEN H., SCHMIDT D., CHEHIMI F. “PhoneTouch: A Technique for Direct Phone Interaction on Surfaces”, UIST'10, New York, USA.
26. KALTENBRUNNER M., JORDA S., GEIGER G. & ALONSO M. “The reacTable*: A Collaborative Musical Instrument”, Proceedings of the 15th IEEE International Workshops on Enabling Technologies:Infrastructure for Collaborative Enterprises (WETICE'06).
27. JORDÀ PUIG S. “Digital Lutherie. Crafting musical computers for new musics’ performance and improvisation”, PhD Thesis (2005), Universitat Pompeu Fabra , Barcelona, Spain.
28. COOK, P. “Principles for Designing Computer Music Controllers”, NIME '01, Seattle, USA.
29. HUNT A., WANDERLEY M., PARRADIS M. “The importance of parameter mapping in electronic instrument design”, NIME '02, Dublin, Ireland.
30. ARFIB D., COUTURIER J-M., KESSOUS L. “Expressiveness and digital musical instrument design”, Journal of New Music Research 2005, Vol. 34, No. 1, pp. 125 – 136.
31. GREENBERG S., BUXTON B. “Usability Evaluation Considered Harmful (Some of the Time)”, CHI '08, Florence, Italy.

32. KUHARA Y., KOBAYASHI D. “Kinetic Particles Synthesizer Using Multi-touch Screen Interface of Mobile Devices”, NIME'11, Oslo, Norway.
33. DERBINSKY N., ESSL G. “Cognitive Architecture in Mobile Music Interactions”, NIME'11, Oslo, Norway.