MUSICAL COMPOSING ON PITCH TIME TABLE BASED TANGIBLE TABLETOP INTERFACE

A THESIS SUBMITTED FOR THE MASTER OF ENGINEERING DEGREE

CONG NING

(B.Eng.(Hons.)) NTU

NATIONAL UNIVERSITY OF SINGAPORE

DEPARTMENT OF SCHOOL OF ELECTRICAL AND COMPUTING ENGINEERING

2010

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Dr Steven Zhou, my project supervisor, for his excellent mentoring and guidance, which have been essential to my accomplishment during the development of the project design. The entire project development would not be such an interesting and fruitful learning and research experience if it were not for Dr Zhou.

I would like to acknowledge Associate Prof Lonce Wyse from Arts & Creativity Lab, not only for his critical comments and useful information related to visual music computations, but also for his patient listening and invaluable advice to my coursework matters. I am grateful for such understanding and help he has offered, which means a lot to me at this stage of my life.

I would like to extend my appreciation to Interactive Multimedia Lab, National University of Singapore, and the staff: Wei Dong, Denial, Yuta and Wei Quan for their time and instructions on the multi-touch table set-ups which facilities to the project.

I would also like to thank the research students from Nanyang Academy of Fine Arts: Jenny Lu, Shan Yan, Liu Wei, Zhang Xiaona and Yin Lian; as well as the research students from National University of Singapore: Zhen Baiyu, Liu Yang, Xu Haibo, Zhang Mo, Bai Jiawei and Xu Chao, for their active participation and their valuable time in the user study experiments.

Last but not least, I am forever indebted to my parents for their unconditional encouragement and continuing spiritual support throughout my endeavor.

LIST OF PUBLICATIONS

ACCEPTED CONFERENCE PAPER

N.Cong, S. Zhou, “The Creative Tangible Musical Interface Design – Interactive Music Creation Platform for Novice Users. ” 7th International Conference on Advances in Computer Entertainment Technology (ACE). 17-19 Nov.2010.

N.Cong, ZY. Zhou, D.Wei “The MusicPatterns - A Creative Tabletop Music Creation Plateform.” International Conference on Educational and Information Technology,(ICEIT) V3-9, T177. 17-19 Sep. 2010.

PENDING CONFERENCE PAPER

N.Cong, ZY. Zhou, Q.Wei, D.Wei “The Study of User’s Behavior of Interactive Actions on Musical Visual Features” 2010 Conference on Design and Architectures for Signal and Image Processing (DASIP) , 26-28 Oct 2010.

N.Cong, ZY. Zhou, D.Wei “The Pitch Time Visualization of Musical Features in Interactive Multi-touch Table Design” International Conference on Exhibition on Computer and Interactive Techinques, 2010.

SUMMARY

This paper introduces the MusicPatterns, a musical tabletop application that enables both novice and musically trained users to compose and play music in a creative and intuitive way. We begin with exposing the demanding value to adopt human computer interaction (HCI) into music creation especially for less musically trained users, and then go into literature review to further explore the current approaches of tangible user interface (TUI) as well as live performance simulation, from where our design philosophy will be stated. Then we will introduce the game design to further illustrate how pitch visual presentations can enable and enhance the user experience of music creation and appreciation. Following that, the project implementation methodologies as well as the major design issues will be discussed with regards to the MusicPattern application.

An early user study on how the users can create and play music on the tabletop application design are also described. We invited 5 students from Nanyang Academy of Fine Arts (NAFA), School of Music Composition Studies, and 6 laymen users from a local university, none of whom are affiliated with this project. The overall participants' reaction was very positive. Users really liked the idea of playing music by visual patterns. The interface was perceived as being very intuitive since the concept of pitch keys and timing does not require much learning to understand. The user study is divided into 3 categories of 6 assignments in total. Detailed observations and discussions on the user study will also be included in the thesis.

With the idea of using visual aids in understanding and operating music features, the applications can be further extended in the future works.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ILLUSTRATIONS

Chapter One Introduction

1.1 Background Study

They say music is the poetry of the air. They say music is the wine that fills the cup of silence. They say music is the medicine of the mind. The say music is the outburst of a soul. By all accounts, there is no single and intercultural universal concept defining what music might be [17]. Every human being, with his/her unique cultural background and life experience, defines music in his/her own manner. However, despite the distinctiveness of the definition possessed by each individual, music is widely regarded as the universal language of mankind. Composers and lyricists create their opuses, which are interpreted by the instrumentalists and vocalists to become music works, which are then appreciated and acclaimed by a massive population of audiences all over the world. To understand why music is able to formulate such a common ground of empathy that resonates amongst people with diverse cultural backgrounds and life experiences, it is necessary to inspect the underlying factor that inspires the creation of music.

Indisputably, it is human emotion that inspires the creation of music. “Music is the most simple; even the primary way of our emotion flow, far beyond our langrage can reach” [2]. The capability of expressing our emotion with sound is rooted in the biological gene of human being, and the similar encryption of audio-to-emotion code that we inherit and share is formed along human history. From the rhythmic howls in the Paleolithic Age, to the symphonies performed by classical orchestras, all the way until nowadays’ experimental works created with the aid of computers, the music representation has evolved to be capable of precisely encoding into sound the complex human emotions: love, hatred, happiness, sadness, gratitude… Concurrently, our seemingly simple and intuitive music listening experience, known as music cognition, has also become vastly intricate and complex. Both processes imply that, with the evolution of music, the sound constituents that compose music have been explosively increased; the elements of the music, which characterize the sound constituents, have been specified in greater details.

Technically speaking, music is an art form with sound as its medium. The sound constituents are governed by elements of the music, which include pitch, rhythm, dynamics, and timbre (these musicological terms will be further explained in the literature review section). Although music is magical in the sense that even the least musically trained people are able to listen and appreciate the musical emotions in one way or another, it is not the only interesting aspect of music just to appreciate it. The experience of music creation is the other (more stimulating) side of musical entertainment that, however, is rather demanding for novice population. Only those with proficiency in mastering the elements of the music are able to assemble them freely to depict their emotions with the creation of music [5].

Similar to learning a language, a crucial part of the music creation involves the ability to transcribe the acoustic form of the music into visual form termed as musical notation, while retaining as much details of the elements of the music as possible. In the domain of music creation, the standard and conventional way is to use text-based notations for music composition and physical instruments for live performance.

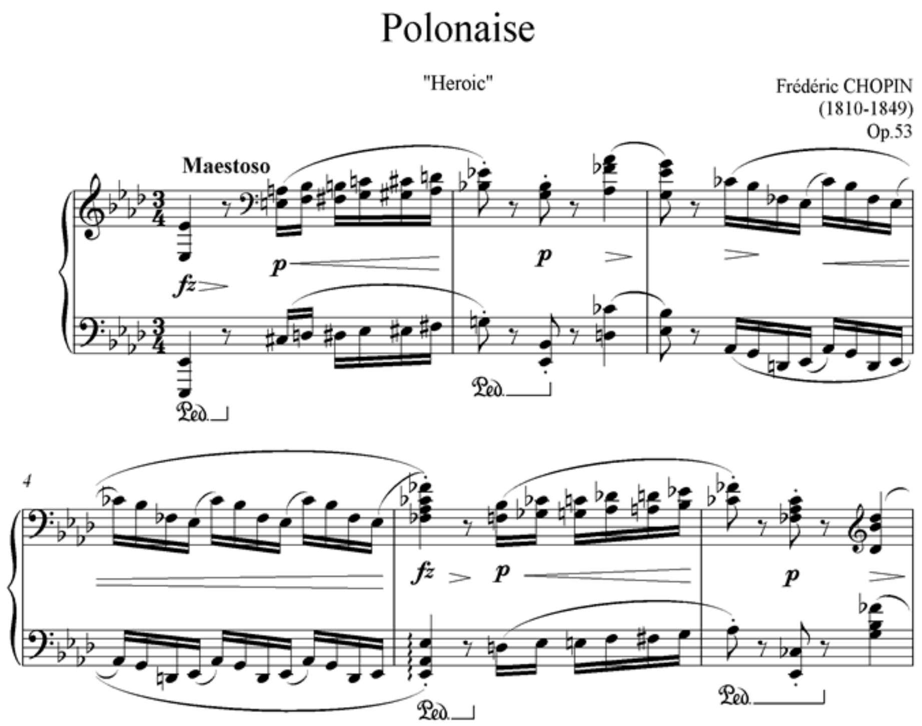


Figure 1.1 An example of modern text-based notation: Polonaise "Heroic"

The modern text-based notation, originated in European classical music, has evolved into a mature system with a very comprehensive collection of symbols and marks, which is adopted by musicians of different genres worldwide. This notation system uses the symbols and marks to represent the various elements of the music to a great extent of detail level, thus it is capable of expressing considerable musical complexity. Besides the precision, the presentation of the text-based notation is comparatively condensed, which was convenient for musicians to make copies of the scores by hand-writing, before machine printing was available. However, the less musically trained people may have learning barriers in text-based notations which is abstract and complicated to understand.

A second aspect of the music creation requires the music to be acoustically reproduced, i.e. the music needs to be performed. Due to the biological restrictions of human voice, the majority of the music works needs to be performed with the aid of some sound-making devices, i.e. musical instruments. It is common that the design of a musical instrument continuously evolves through a prolonged span of time to expand the capacity of articulating richer elements of the music, and the mechanism of the instrument also becomes more complex. As a result, it may take a long time of musical instrument training and practicing to even master a single instrument in live performing. Because of the two aforementioned aspects, the music creation has kept distance from laymen and only opens to virtuosos.

The introduction of the computer has significantly changed the horizon of music creation. Not only does the computer technology serve the professional musicians in their composition and performance, it also offers an alternative way of music creation besides the conventional methods that is more accessible for laymen. Aiming to lower the bar to music composing and performing, and to let novice to music to enjoy the essence of music creation with less effort, researchers have developed numerous application platforms introducing the Human Computer Interaction (HCI) technology into the music creation, which has indeed opened a new dimension in term of enriching user experience in music creation. Comparing to the conventional way of musical creation, many HCI applications adopt tangible user interfaces (TUI) [1] to replace the dreary text-based musical notation with visual graphical representations, and at the same time, simulate a live music playing environment that is easier in operation than the real classical musical instruments. Moreover, the computer-aided platform has a digital nature, contributing to the merits of ease of editing, reload ability, multi-track, multi-tone, which support remote and collaborative works.

However, the HCI simulated musical creation platforms have also limitations with comparison to the conventional music composition and performances. First, almost all TUI-based application platforms have limitation in their physical dimensions that restricts the length of music piece they can contain. Second, the various graphical musical notations adopted by these applications, although intuitive and easy to understand, are inferior to the traditional text-based notation in preserving the details of the elements of the music. Furthermore, with the design objective to simulate current existing musical instruments, an HCI system will always be considered less sophisticated than a real musical instrument in terms of the sound articulation mechanism, thus it can never totally achieve what a real instrument can do.

The trade-off is obvious. The “professional solution” (i.e. conventional way of music creation with text-based notation and real musical instrument) excels in the preciseness and the abundance in musical details, as well as the limitless form of presentation, whereas the “layman solution” (i.e. HCI simulated musical creation application platform with visual-based notation and TUI) excels in the intuitiveness and the ease of hands-on. The solution now lies in between: where is the equilibrium point? With that question in mind, we are aiming to inject an intuitive and easy concept into the design of a creative composing/ performing integrated HCI platform which has the potential of usage in musical education and entertainment fields.

1.2 Schedule of the project

A Gantt chart is shown below depicting the task allocation along the span of the project period.

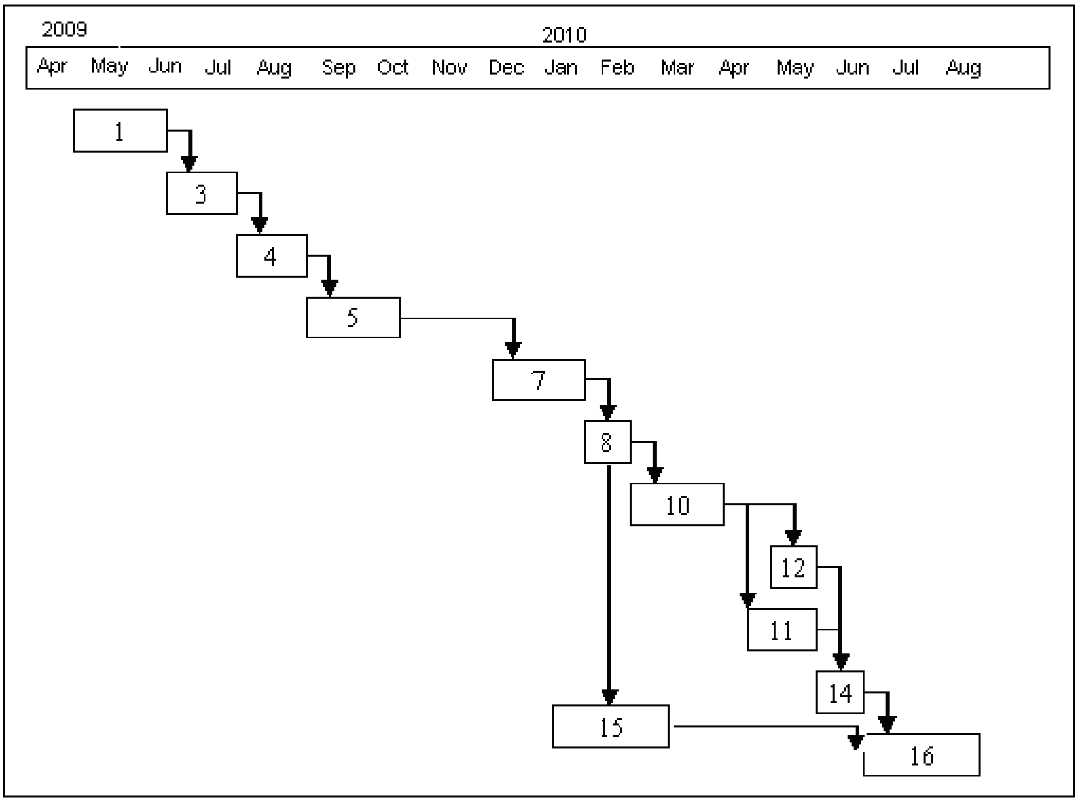


Figure 1.2 Gantt chart of the project schedule

No.	**Date**	**Activity**
1	mid April 09 – late May 09	Browsing through materials and previous works related to the project.
2	01 Jun 09	First discussion with supervisor to identify project goals.
3	mid Jun 09 – mid Aug 09	Analysis of the pros and cons of the previous music tabletop applications and setting design objective.
4	Mid – late Aug 09	Initial implementation of the pitch time table layout design.
5	1st Sep 09	First meeting with Professor Lonce Wise to discuss about the table layout idea; Starting to add in pitch time table musical features into the application.
6	Nov 09	Design of the stocking tube mode

7	Early Dec 09 – Jan 10	Testing and result analysis for the implementations issues. Implementation of the demo application GUI.
8	Feb 10	Review and summarization of project work done. Preparation of the interim report.
9	25 Jan 10	Submission of the interim report.
10	Mar 10 – early April 10	Implementation and integration of the small keyboard tube mode.
11	mid April 10	Preparation of 1st draft of the thesis.
12	starting May 10	Modification of the structure of the program developed so far. Writing the Sig-graph and ICEIT conference papers.
13	2^{nd} Jun 10	ICEIT paper notification of acceptance.
14	Late Jun 10	Revision and amendment of the final thesis paper.
15	9 Apr 10 – early Jul 10	User study
16	Jul 10 – mid Aug 10	Modification of the final thesis paper and documentation of the codes.
17	22 Aug 10	Oral presentation and thesis submission.

Table 1.1 the Project Schedule

1.3 Scope of the thesis

In this report, the design and development of the tangible musical interface will be discussed in great detail. Significant problems encountered along the implementation process will be described together with their solutions. In several scenarios, alternatives of design techniques available will be compared, and the reasoning will be elaborated for the choice of the final technique adopted.

Testing procedures as well as result manipulation and presentation will also be covered. They include the testing process planning and execution, result tabulation and analysis, and the design and implementation of the demo application GUI.

Proposed but as yet unimplemented works will be mentioned but will not be explained in detail in this report. Examples of such future works are human voice recognition as the input over finger operations and further editing musical tool box designs.

1.4 Organization of the thesis

The rest of the thesis is organised as follows. Chapter Two gives a literature review on the TUI tabletop hardware setups and then continues to explore the pros and cons of current existing tabletop musical applications, followed by a briefly introduction of various issues related to musical features that will be used in this project design. An overview of the game design logic and game design interface is presented in Chapter Three. Chapters Four addresses the implementation details of software data flows as well as the difficult design issues encountered during the software design. Chapter Five describes the testing procedures and the early user study results for the MusicPattern platform with detailed analysis. The design and implementation of the demo application GUI is covered in this chapter too. Chapter Six concludes the entire project with the mentioning of future works related to the topic.

Chapter Two Literature Review

2.1 Multi-Touch Tabletop System

Touch screen technology has a long-standing history with one of the first known multi-touch system developed around 1982-83 by the Input Research Group of the University of Toronto under the leadership of Bill Buxton. It became widely popular by the release of Apple's iPhone in 2007. Various technologies are used for the implementation of multi-touch. The three most common ones are (a) capacitive sensing, e.g. iPhone; (b) resistive sensing e.g. Touchpad; (c) computer vision systems e.g. Microsoft Surface.

The most prominent advantage for finger-touch supported system is that the interaction with an application through directly touching the graphical elements is more "natural" or "compelling" than working with an indirect pointing device, typically mouse. Previous research has pointed out that, for multiple-user collaborative tasks, the finger-touch tabletop is preferred over mouse input, but other factors such as insufficient accuracy due to ambiguous touch area and accidental input by other fingers can affect the preference greatly. By enhancing the "natural" and "compelling" features of direct finger-touch, multi-finger gestures are introduced to the tabletop system. Since it is no longer one single cursor moving around the interface and delivering control, we can exploit the multi-touch features to design multi-finger gestures. The designer of the system must address questions such as how the gestures should be mapped to the various functions. Previous research has introduced the concept of gesture registration, gesture relaxation and reusability, which are generally based on the life cycle of the gestures. Overall, the challenges raised in multi-finger gestures are mainly pertaining to functional match and implementation of gestures. While finger gesture is one way to trigger functions, finger orientation can also be interpreted as control information. Research on finger orientation has shown that it is a feasible and valuable input dimension that can be utilized for novel interactive interface. This enlightens the researchers to explore further in the domain of finger-touch input.

These characteristics highlight that users doing collaborative task on the tabletop may interact but should not interfere with each other, and the device they use should be easy to maintain, cost efficient, and stand-alone. There have been various designs proposed in previous research outstanding for different characteristics aforementioned. Thus, when employing different design to construct a tabletop system, it is essential to determine what characteristics would help our application on the tabletop system so as to achieve its best utility.

The collaborative consideration also applies not only to hardware setup but also to software design. “Fluid interaction and meaningful visualization is the key for multi-surface, multi-device, interactive spaces to become the users' true cognitive prosthesis.” [20] The interface presented to users should provide a convenient space for them to share and organize information, and seamlessly integrate multiple users into the collaborative task.

The hardware is making use of vision (camera) based multi-touch sensing. Camera based sensing itself can be implemented in different ways. Some of the examples are FTIR (Frustrated Total Internal Reflection), DI (Diffused Illumination) and LLP (Laser Light Plane).

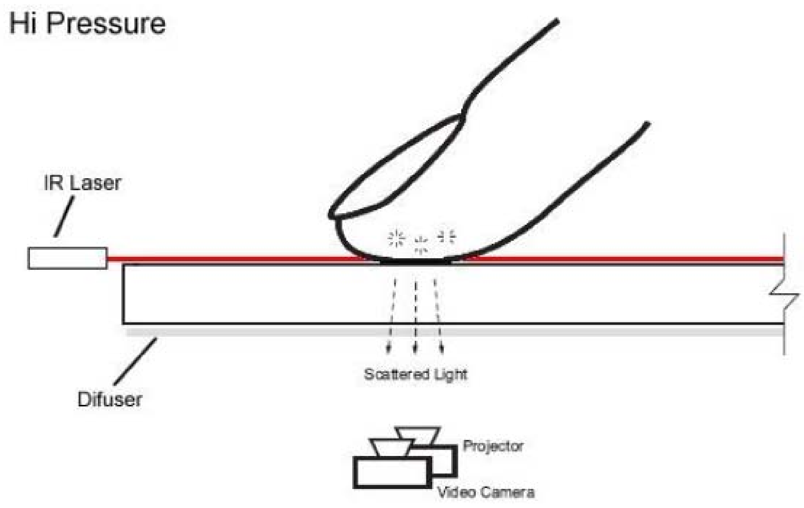


Figure2.1: interactive tabletop set-ups

For the hardware setup of a laser light plane, four infrared(IR) lasers are installed at the four corners of the touch table to make an even laser light plane about 1mm in thickness exactly above the table surface. The IR lasers used are of 850 nm wavelength, 5V and 28mW of power. The higher the power of the IR laser, the brighter the laser lights it generates. Line lenses are used to convert the point light from the lasers into a plane. Looking directly into the IR laser lights may cause serious injury to the eyes, so the users have to be aware of the hazard and avoid such situations in the experimental setup.

A camera and a projector are placed behind the surface of the table. When there is a touch on the surface, the laser light is reflected and captured by the camera as a blob. The camera used for this purpose should be IR sensitive to capture the reflected IR light. Most of the normal digital cameras are IR sensitive, but use an IR filter to block off those higher wavelength spectra to allow only visible light to be captured. A common approach to set up multi-touch tables is to buy a normal webcam with good fps (frames per second) rate and manually replace the IR filter with a negative film or some other filters to block visible light. The hardware setup for this project employs an industrial camera “Firefly” from the company Pointgrey which is already IR sensitive with a visible light filter. For projectors, short throw rear projectors will be ideal and easy to use. A normal projector can also be used by projecting onto a mirror and reflecting to the table surface. This makes the projected rays to travel far enough to converge nicely on the table-top and at the same time inverts the image for back projection. The surface of the table is made of transparent acrylic so that the reflected IR light can pass through it to the camera underneath. A diffusive layer is added below the acrylic as a surface for the projection.

2.2 Multi-Touch Tabletop Protocols

There are different processes and various protocols for a touch sensed at the output screen to be recognized as an input gesture in the particular ‘multi-touchable’ application that is running on the screen.

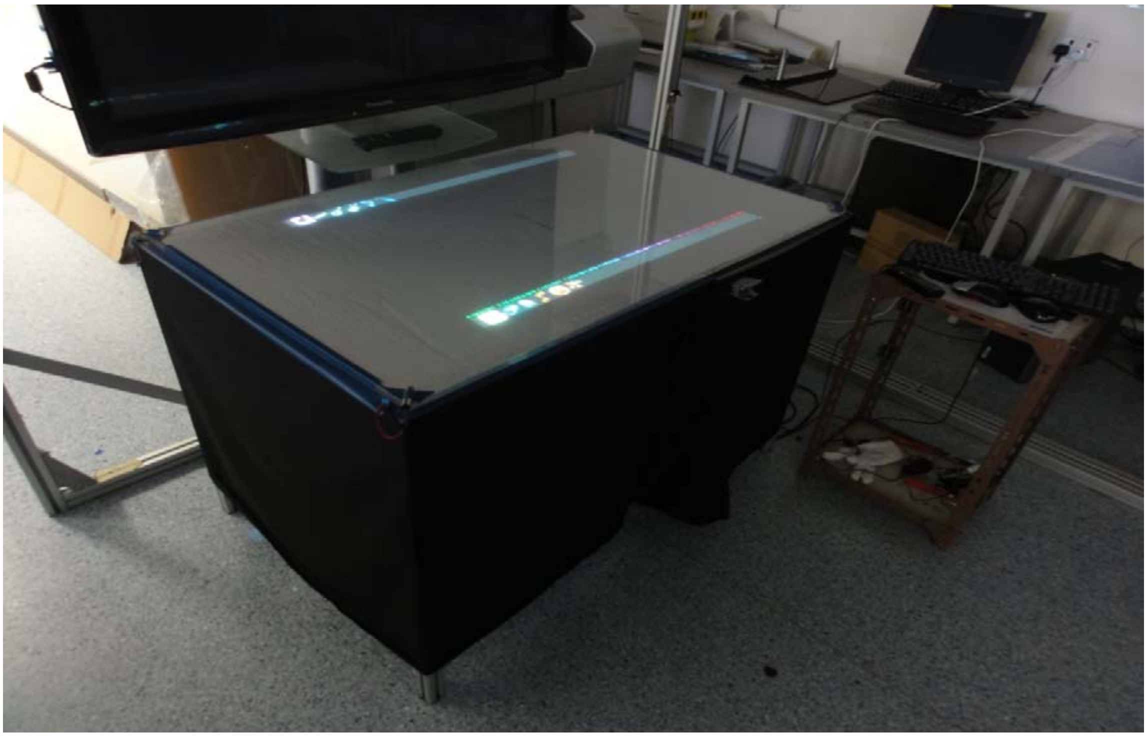


Figure2.2: The hardware setup used in this project design

TUIO (Tangible User Interface Object) [15] protocol developed by the 'Reactable' is used for communicating the tracked blob data on the table top to the computer. This protocol has options to communicate the position, size and relative velocity of the blobs tracked. TUIO is built based on OSC (Open Sound Control) protocol which in turn is built based on UDP. TUIO is a standard protocol to send OSC messages for table-top touch recognition. TUIO defines a robust computer vision engine and a communication interface between table-top controller interfaces including hand (and finger) gestures and underlying applications. The objects (in the case of fiducial markers) and touch (when using hand or finger) are detected by a sensor system and are identified for their position and orientation (only for objects) on the table-top. Each object or touch gets a unique 'session ID' and a non-unique ID called 'objects ID' (objects) or 'cursor ID' (touch).

"Open Sound Control (OSC) [21] is a protocol for communication among computers, sound synthesizers, and other multimedia devices that is optimized for modern networking technology." OSC [22], originally developed by the Center for New Music and Audio Technology of UC Berkeley, is a network communication protocol built on UDP and is successfully implemented in many hardware devices for sensor measurement. Advantages of OSC include speed, precision, rich type support and human readability. An OSC message has a name/value pair (a 'Dictionary' data type which uses a hash table) and an optional time tag. The time tags are of high resolution for the accuracy. Rather than sending messages one by one, OSC can send bundles of messages which may occur and have to be processed simultaneously.

A tracker is a "program which takes images from a camera, puts them through several filters, and finally reports the position, size, and relative movement of blobs over some protocol." Two of the most popular open source trackers are 'Community Core Vision (CCV)' [23] (older versions known as tBeta) and ‘Touchlib’[24]. These frameworks serve as a solution for computer vision and machine sensing. They take in a video stream as input and output tracking data such as blob coordinates and events such as finger down, moved, released etc. This project setup uses CCV, which is the most recent and more stable based on user experience.

The tangible tabletop system supports the visual display with real-time touch-screen feedbacks, thus establishing a two-dimensional tangible interface linking real world objects or fingers to the computer based virtual graphics. Besides the visual graphics, the computer system is also capable of generating the real-time audio feedback with respect to the object actions. The advantage of developing musical applications on top of tabletop system is that the direct object operations can have both real-time visual and audio response [31]. With proper designed software program, the tabletop system has the potential to present a new way of playing and editing music, which enables users to touch, see, and hear simultaneously, and makes learning and creating music more intuitively and comprehensibly.

2.3 Table-top musical applications

ReacTable (2004)

Figure 2.3: The reacTable interface [25]

The reacTable [8] is a multi-user electro-acoustic musical instrument with a tangible tabletop user interface which allows expressive collaborative live performances for professional musicians. The reacTable can be played by manipulating a set of objects known as fiducials that are distributed on top of a touch-sensing table surface. In addition to the sound which is produced by manipulating the fiducials, the reacTable also provides additional visual feedback [10]. This feedback projects a visualization of the sound-flow onto the table surface which is shown to the users as colorful sinusoidal waves or distorted lines as shown in Figure 2.3.

The framework of reacTable, the reacTIVision , has also been used by interest groups to create music table. MixiTUI and AudioTouch are examples of multi-touch table which uses reacTIVision. The reacTable supports remote collaboration where users can perform concurrently at different locations on the same piece of song. Demonstration of such remote collaboration was performed over the internet between Spain and Austria. The reacTable is targeted more towards professional musicians for concert purposes rather than beginners. Each fiducial carries different function or sound effect , and this might require knowledge to control them, thus posing a potential obstacle for beginners to maneuver. It is discovered that a lack of physical control over the fiducials may occur if too many fiducial blocks are placed on the table. Furthermore, having too many fiducials blocks may create problems in storage and managing in the long run. There is little emphasis on music theory as each fiducial carries an sound effect but not pitch tones. Music composition is done in regular intervals, e.g. regular beat pattern. Therefore, it is tedious to compose music in a continuous format.

Scrapple (2005)

Figure 2.4: User manipulating spectrograms on Scrapple [27]

Scrapple is a camera-based spectrographic performance instrument with a tangible interface which uses an ‘augmented reality’ (AR) [28] overlay to provide the user with *in-situ* visual feedback regarding the state of the system. The tangible objects are known as spectrograms, or diagrams which depict the frequency content of sound over time. It is noted that Scrapple is not a touch sensing table. Users can arrange a variety of dark rubber and felt objects on a three-meter-long table, where the positions of the individual elements on the horizontal and vertical lines determine the rhythms and pitches. The three-meter-long musical score is scanned at regular intervals and the music notation created by the user is played back repeatedly. The Scrapple uses a hybrid synthesizer (combining a granular and an additive sound synthesizer) which produces four-second-long audio loops. Video projections are installed on top of the table to refine the objects arrayed upon the table, thus achieving accurate spectral synthesis and image registration.

The system implementation is targeted with respect to rhythm and pitch. It follows the music theory format, where the y-axis determines the pitch and the x-axis determines the rhythm. This allows the user to understand and visualize music and sound by using shapes. However, the three-meter-long table used for the Scrapple system imposes a space constraint which is not desired by users, and it will become messy in the long run with too many tangible objects to keep track. In the case of multiple users working collaboratively on the same table, it is difficult to control the markers if too many of them are used at the same time on the touch table. Furthermore, music is composed in periodic timeline; therefore it is not allowed to have composition of a proper song with verse, chorus, and bridge in a continuous structure.

AudioTouch (2008)

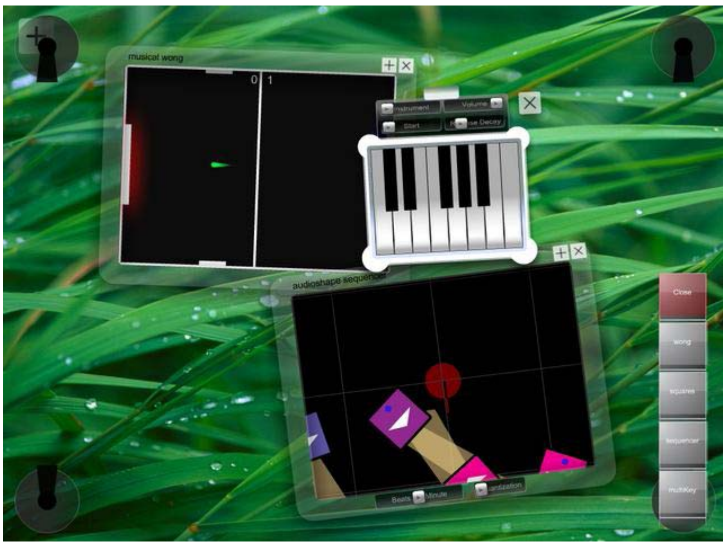


Figure 2.5: Interface of AudioTouch [30]

AudioTouch is an interactive multi-touch interface for computer music exploration and collaboration. The interface features audio based applications that allow multiple users to generate simultaneous touch inputs on a single multi-touch screen. The goal of the interface is to achieve user interaction with the technology (specifically music based) in a natural way. The key aspect of the design is a natural user interface, where users can interact through gestures and movements, while directly manipulating musical objects. By using a fully touch-based table, it removes the need of having tangible objects.

mixiTUI (2009)

Figure 2.6: Users collaborating locally on mixiTouch [33]

mixiTUI [9] is a tangible sequencer that allows electronic musicians to import and perform electronic music. mixiTUI is developed in collaboration with electronic musicians, with a focus on live arranging, on visualizations of music, on tokens that represent key elements in live performances, and on how the audience experiences the tangible interface. It is built on the existing framework, the reacTIVision[26], which supports the basic architecture of a music table.

One major improvement over the reacTable is that there is a session control bar at the side of the screen which provides control over a group of fiducials known as tokens. For example, two types of different drum beats belonging to the loop token category can be muted at the same time once the user shifts the control away at the session control bar. This implementation solves the problem of managing multiple fiducials at the same time in the reacTable.

Similarly to reacTable, the mixiTUI is targeted more towards professional musicians for concert purposes than beginners. Each tokens carries different function or sound effect which may requires knowledge to control them and this may pose an obstacle for beginners to maneuver.

There is little emphasis on music theory as each token carries an effect and not pitch tones. In addition, music composition is done in regular intervals, e.g. regular beat pattern. Therefore, it is tedious to compose music in a continuous format.

Summary of Evaluated Works and Projects

Table	Year	Pros	Cons
reacTable	2004	• Supports remote collaboration.	• Not suitable for beginners • Lack of control over multiple fiducials • Little emphasis on music theory • Does not support music composition in a continuous timeline. Uses loop-back.
Scrapple	2005	• Follows the music theory, where y-axis determines the pitch and the x-axis determines the rhythm. • Provides visualization of the frequency of sound in a form of shapes.	• Difficult to control if too many spectrograms are used at the same time. • Does not support music composition in a continuous timeline. Uses loop-back.
AudioTouch	2008	• A fully touch-sensing table • does not require tangible objects.	• Little emphasis on music theory. • Uses game application to stimulate sound and not music.
mixiTouch	2009	• Allows control over a group of fiducials at the same time. • Supports remote collaboration.	• Not suitable for beginners • Little emphasis on music theory • Does not support music composition in a continuous timeline. Uses loop-back.

Table 2.1: Comparison of reviewed Multi-touch Tables

From Table 2.1, it is observed that most of the touch-tables except AudioTouch still use tangible objects to control the software application. This indicates that there is still an inclination to use tangible objects. However, the complication of managing the objects physically [6] is an obvious shortfall for using tangible objects.

It is also observed that since the reacTable in 2004, slight improvements have been made to the subsequent products. mixiTouch (2009) solves the problem of managing multiple fiducials or tokens at the same time by implementing a session bar which can control a group of tokens easily. However, it is noted that mixiTUI still uses the existing framework of reacTable which is the reacTIVision, and hence there is little emphasis on music education. mixiTUI still uses effects and sounds but not pitch tones.

Scrapple (2005) views music creation in a traditional music theory context. However, music composition is done in a play back or a loop-back manner. It does not allow composition of a proper song with verse, chorus, and bridge in a continuous structure.

2.4 Music Features Review

In music theory, common fundamental parameters can be distilled and identified from music of various styles, genres, and historical periods. As mentioned in the introduction, these parameters or elements of the music include, but are not limited to, pitch, rhythm, dynamics, timbre, etc. The articulation of each particular sound can be described with respect to all the elements, i.e. the elements characterize the sound in different dimensions. It is necessary to study these elements as they are the basic ingredients which have been used in the music creation, as well as the essential components that need to be designed for the computer-based music applications.

Rhythm

Rhythm is the parameter that describes music in terms of timing factors. Technically speaking, rhythm is the arrangement of sounds and silences in time. In a general sense, the duration of a piece of music is measured in number of bars. The term bar is derived from musical notation, which represents a segment of time defined as a given number of beats of a given duration.

When we listen to music, the steady underlying pulses to which one could clap are called the beats. Among the beats, there are on-beats which are more stressed, and there are off-beats which are less stressed. The arrangement of on- and off-beats is repeated periodically, which divides the music into equal segments. These segments that we feel can be regarded as the acoustic definition of the term bar.

For the articulation of a single sound, or the music note, its length in time can be changed by varying attack, sustain and decay characteristics of the note. The duration of the music note from attack (onset) to decay (offset) is called the note value. The table below lists the most frequently used note values, sorted in descending order by its duration. The duration of the silence or rest is also defined in the same way.

Note	**Rest**	**American name**	**British name**
𝆷		longa	longa
𝅜	𝄺	double whole note	breve
𝅝	𝄻	whole note	semibreve
𝅗𝅥	𝄼	half note	minim
𝅘𝅥	𝄽 or 𝄿	quarter note	crotchet
𝅘𝅥𝅮	𝄾	eighth note	quaver

𝅘𝅥𝅯	𝄿	sixteenth note	semiquaver

Table 2.2: Musical Notation of the Note Values

In the above table, the length of each note value is equal to twice the length of the note value immediately below it, in other words, the shorter note values are derived by slicing the length of the note value immediately above it in half. The American naming convention of the note values also gives a hint to this feature. It can be illustrated in a hierarchical form, as depicted in the figure below.

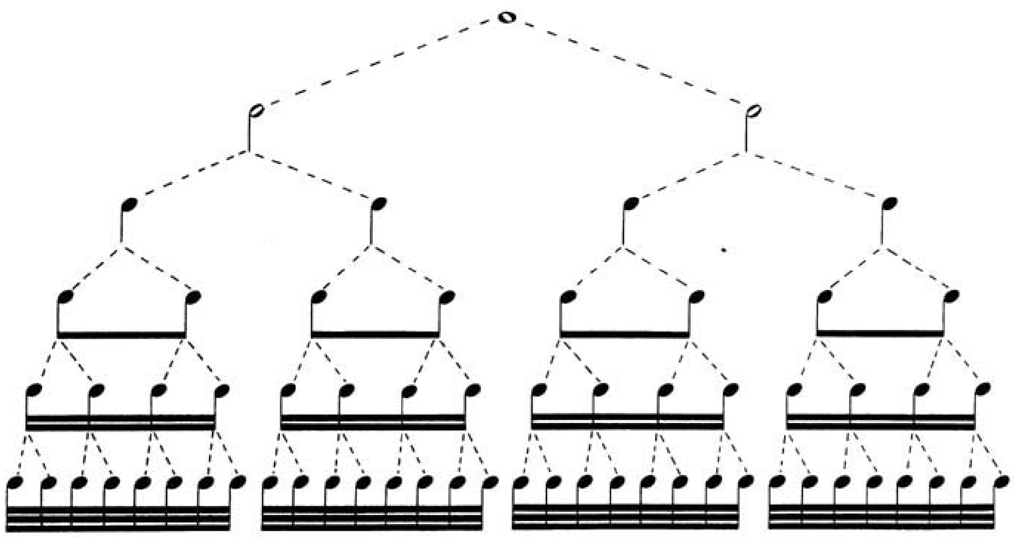


Illustration 2.1: Note Value Hierarchy [46]

Time signature, or meter, is defined as the number of beats (in terms of music notes) per bar in a music piece. For example, a 2/4 meter indicates two crotchet beats (quarter notes) in each bar, and a 6/8 meter indicates six quaver beats (eighth notes) in a bar. The tempo is a quantitative measure of the frequency of the beats in the unit of BPM (beats per minutes). At a 2/4 meter, the tempo is the number of crotchet beats per minute, whereas at 6/8 meter, it is the number of quaver beats per minute.

Pitch

Pitch represents human perception of audio frequencies of a sound. [18] Although it is a mere psychological perception to most human beings as "higher" and "lower", we are able to borrow the scientific measure of frequency to quantify it. Normally A4=440Hz is used as the reference pitch in concerts and is the American standard pitch. Based on this reference pitch, the fundamental frequencies of the 12 pitch class notes with their octave alignments are derived as shown in the Table below.

Sub-band No		01		02	03	04	05	06	07	08
Octave scale		~ B1	C2 ~ B2	C3 ~ B3	C4 ~ B4	C5 ~ B5	C6 ~ B6	C7 ~ B7	C8 ~ B8	All the higher octaves in the (8192 ~ 22050) Hz frequency range
Freq-range (Hz)			64~128	128~256	256~512	512~1024	1024~2048	2048~4096	4096~8192	
12 Pitch class notes	C	0 ~ 64	65.406	130.813	261.626	523.251	1046.502	2093.004	4186.008	
	C#		69.296	138.591	277.183	554.365	1108.730	2217.460	4434.920	
	D		73.416	146.832	293.665	587.330	1174.659	2349.318	4698.636	
	D#		77.782	155.563	311.127	622.254	1244.508	2489.016	4978.032	
	E		82.407	164.814	329.628	659.255	1318.510	2637.02	5274.04	
	F		87.307	174.614	349.228	698.456	1396.913	2793.826	5587.652	
	F#		92.499	184.997	369.994	739.989	1479.978	2959.956	5919.912	
	G		97.999	195.998	391.995	783.991	1567.982	3135.964	6271.928	
	G#		103.826	207.652	415.305	830.609	1661.219	3322.438	6644.876	
	A		110.000	220.000	440.00	880.000	1760.000	3520.000	7040.000	
	A#		116.541	233.082	466.164	932.328	1864.655	3729.310	7458.62	
	B		123.471	246.942	493.883	987.767	1975.533	3951.066	7902.132	

ISO 16 standard specifies A4 = 440Hz and it is called as concert pitch

Table 2.3: Music note frequencies (F0) and their placement in the Octave scale sub bands.

Note that the frequency is a continuous measure, whereas the traditional musical pitch is a discrete subset of the sound frequency spectrum. In real life scenario, the music pitches articulated by musical instruments often fluctuates due to the physical condition of the instruments. However, human perception tolerates such deviation of frequency and is still capable of recognizing the pitch so long as the deviation is within certain tolerance range.

When a sequence of pitches is arranged in sequence, a melody is created. Melody is monophonic in nature. By simultaneously playing multiple pitches or sequences of pitches, we create harmony, which is polyphonic in nature. The simplest harmony is produced by playing a chord, which refers to the set of music notes played simultaneously. The difference of the intervals between the pitch frequencies of the notes that constitute a chord affects our perception of the "chromatics" of that chord. A major chord, for example, is considered consonant, stable, and bright, whereas a minor chord sounds relatively darker than the major chord. Both augmented and diminished chords introduce a suspense feeling, while the diminished chord sounds more perturbed than the augmented chord. If we arrange the twelve pitch class notes in the octaves into a circumference that resembles an analogue clock face, and depict the notes of the chord as the clock hands, we create a pitch constellation of the chord. The figure below is an example of the pitch constellation of triadic chords (chords that is constituted by three distinct notes) in the key of C. Note that the chromatics of the chord are not relevant to the absolute frequency of the pitch, but the relative "distance" or interval between the constituent pitches along the frequency spectrum. In other words, for the same chord in different keys, the pattern is the same on the pitch constellation.

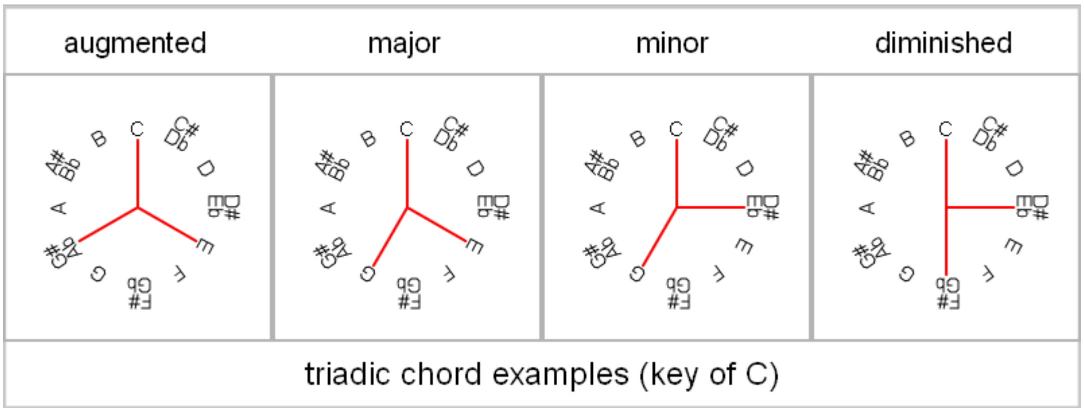


Figure 2.7 Pitch Constellations of Triadic Chords (Key of C) [46]

Dynamics

Dynamics refers to the volume of a sound or note, which transcribes into the amplitude of the sound in technical sense. In the scope of an entire piece of music, dynamics can also refer to various stylistic or functional (velocity) aspect of the rendering of the work. In conventional musical notation, dynamics are not indicated by specific volume levels, but is rather denoted as the relative loudness among the constituent notes in the entire work [7], sometimes even with reference to the ensemble as a whole if played in an orchestra.

Timbre

Timbre refers to the quality of sound that distinguishes the various mechanisms of sound productions, including musical instruments and human voices. The same note with identical pitch, duration, and volume, being produced by a piano and a violin, is perceived differently by us. In acoustic terms, the distinction in timbre can be attributed to the physical properties of the sound-producing components of various musical instruments or the vocal chords of individual human beings, which affects the waveform of the sound they make.

Comprehensive acoustic analysis is required to have an in-depth understanding of the timbre. Here we only list out the five major acoustic parameters to determine the "elusive attributes of timbre" as identified by J.F. Schouten [18]:

1. The range between tonal and noise-like character.

2. The spectral envelope.

3. The time envelope in terms of rise, duration, and decay (ADSR — attack, decay, sustain, release).

4. The changes both of spectral envelope (formant-glide) and fundamental frequency (micro-intonation).

5. The prefix, an onset of a sound quite dissimilar to the ensuing lasting vibration.

Music Scale

A sequence of notes with their pitches arranged in ascending or descending order, which forms a particular context that can conveniently represent musical works, is called a music scale. The distance between two successive nodes in a scale is called a scale step. The most commonly known music scale is the diatonic scale, which consists of seven basic notes represented by successive alphabetical characters C, D, E, F, G, A, and B. These basic notes can be produced by depressing the white keys on the piano keyboard. By adding the notes produced by depressing the black keys on the keyboard into the diatonic scale, a chromatic scale is generated. The chromatic scale consists of twelve notes, namely C, C♯(D♭), D, D♯(E♭), E, F, F♯(G♭), G, G♯(A♭), A, A♯(B♭), and B. The twelve pitches of the chromatic scale are equally spaced, each a semitone apart. Following the note B in both scales, by adding another scale step, we reach a higher C note with double the frequency of the C that we start with, and the scale repeats itself from that higher C. The interval between the lower C and the higher C with double the frequency is termed as an octave. Like most of scales, both diatonic and chromatic scales are octave-repeating, meaning the pattern of the notes is the same in every octave, and repeats in a cyclic manner.

By filtering out the absolute pitch frequency and focusing on the relative distance between adjacent notes, it is obvious that the pitch progression in a music scale from one note to another is either a half step (known as a Semitone, abbr. S) or a whole step (known as a Tone, abbr. T). The first note in the scale is known as the tonic, and is so-called the tone-note from which the scale takes the name. Starting from the tone-note, we can depict the scale by specifying whether the interval between two successive notes is an S or a T. For instance, a major diatonic scale can be transcribed as T-T-S-T-T-T-S. Therefore, by fixating the tone-note at certain pitch (e.g. G♯) and applying the pattern along the note sequence, we define a major scale with the name of tone-note (e.g. G♯ Major). The figure below illustrates the generation of all major scales with the same interval pattern.

MAJOR SCALES	T	T	S	T	T	T	S	
C Major	C	D	E	F	G	A	B	C
G Major	G	A	B	C	D	E	F#	G
D Major	D	E	F#	G	A	B	C#	D
A Major	A	B	C#	D	E	F#	G#	A
E Major	E	F#	G#	A	B	C#	D#	E
B Major	B	C#	D#	E	F#	G#	A#	B
F# Major	F#	G#	A#	B	C#	D#	E#	F
C# Major	C#	D#	E#	F#	G#	A#	B#	C#
F Major	F	G	A	Bb	C	D	E	F
Bb Major	Bb	C	D	Eb	F	G	A	Bb
Eb Major	Eb	F	G	Ab	Bb	C	D	Eb
Ab Major	Ab	Bb	C	Db	Eb	F	G	Ab
Db Major	Db	Eb	F	Gb	Ab	Bb	C	Db
Gb Major	Gb	Ab	Bb	Cb	Db	Eb	F	Gb
Cb Major	Cb	Db	Eb	Fb	Gb	Ab	Bb	Cb

Figure 2.8: Major Scales of Various Tone-Notes

Chapter Three Game Design

3.1 The Game Design Philosophy

The main design objective is to provide a tangible interface for both novice and musically trained users to visualize the music features and to offer an intuitive music creation platform to them. The first challenge is how to present music on the two-dimensional tabletop surface. As the objective of this tabletop application is to use direct finger operations to play and edit music, the user interface design must have a visual appearance that is easy enough to understand, and more importantly, it must be able to further interpret finger actions into meaningful operations on musical features [32]. In order to do that, we first go back to the definitions of key musical features reviewed in Chapter Two. As introduced in section 2.4 Music Features Review, the most important musical features are Rhythm, Pitch, Dynamics, Timbre, and Musical Scales, and this can be further interpreted as four dimensions of information:

1. Pitch, referring to perceived fundamental frequency of a sound.
2. Timing, including how pitch will flow in the time domain and how the continuously varied pitches form a Pitch-Flow (abbreviated as PF [34] in the following context) in each time domain.
3. Volume, referring to the volume of any certain pitch at any time instance.
4. Tones, referring to the musical instrument timbre of a certain PF.

However, notice that the combination of the above four elements does not necessarily form music [43]. There are other advanced features built on top of this four elements[35] such as Chorus, Rhythm, Mode, Air, Allegro, Overtone, as well as the musical terminologies such as Overture, Solo, Sonata, Syncopation and so on, and even with all these advanced music features cannot define the scope of music. This design does not go into the discussion of what is music, and focuses more on the design of adopting Pitch, Time, Volume and Tones as primary dimensions to provide an audio-visual user interface for music creation and editing.

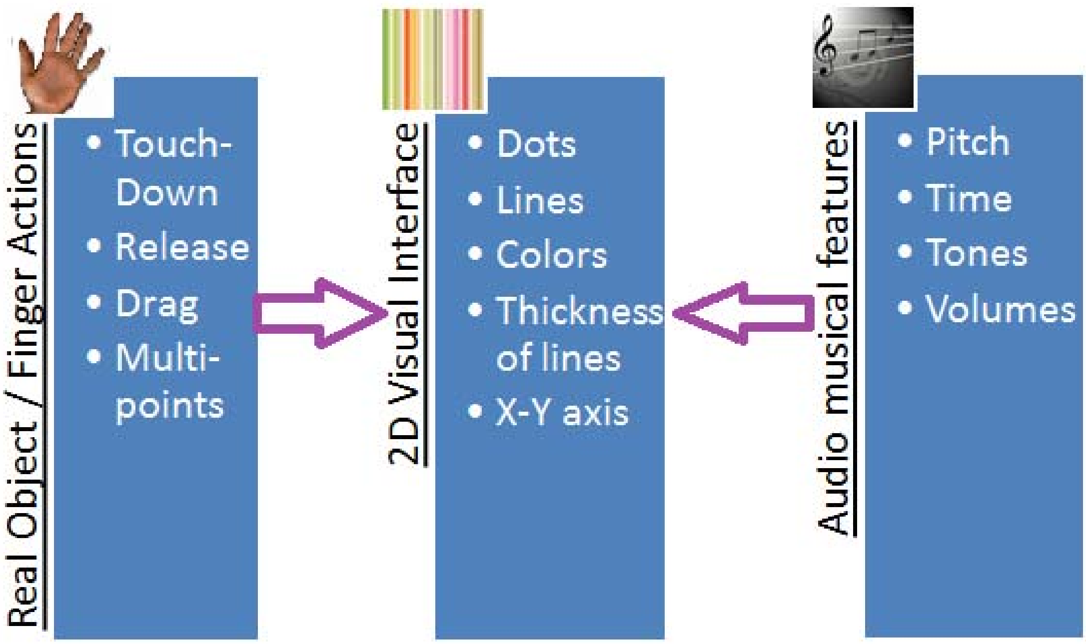


Figure3.1. Object/Visual/Audio relation

Figure 3.1 shows the relationship between the real objects, the elements of the tangible interface and the features of audio music. It can be observed that the two-dimensional tabletop visual graphic interface serves as the bridge that connects the finger operations to music features. A successful tangible tabletop musical application should establish an intuitive and feasible linkage between real objects to visual features as well as linkage between audio musical features to visual features [36]. First of all, we explain how to set up the link between 2D visual interface and audio music features. To map the four-dimensional information of Pitch, Time, Volume and Tones onto the two-dimensional tabletop interface, we introduce two extra visual dimensional features namely Color and Thickness. The exact mapping relation is as follow:

Pitch map to x-axis of the 2D canvas;

Time map to y-axis of the 2D canvas;

Tones map to the Color of the PFs [37];

Volume map to the thickness of PFs;

Assuming that x-axis scales from the left to the right are mapped to pitches from low to high. If a certain single pitch starts to voice, this can be represented as a dot shown below:

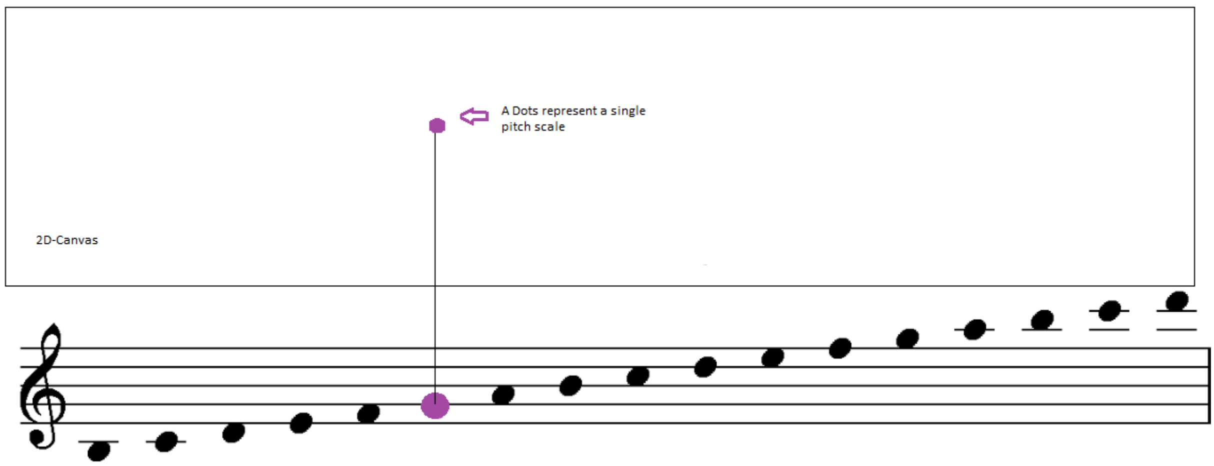


Illustration 3.1: Single point pitch

With the pitch flow from t=t1 to t2, this forms a PF, represented by a line on 2D tabletop.

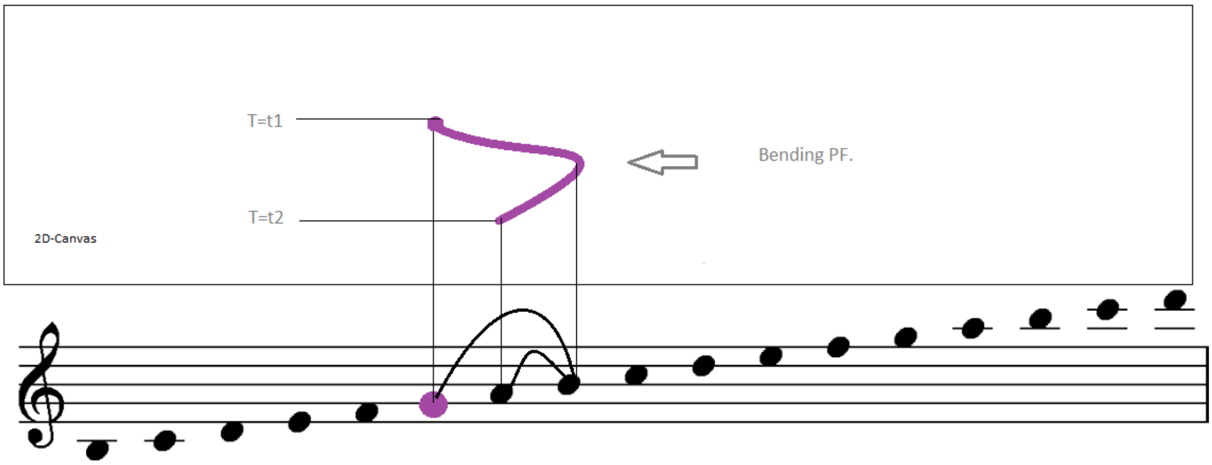


Illustration 3.2: Single pitch flow

The sample PF in Illustration 3.2 is perceived as a bending pitch starts at T=t1 and ends at T=t2, and it crosses up and down over three pitch scales. Next, we add another instrument represented in green color:

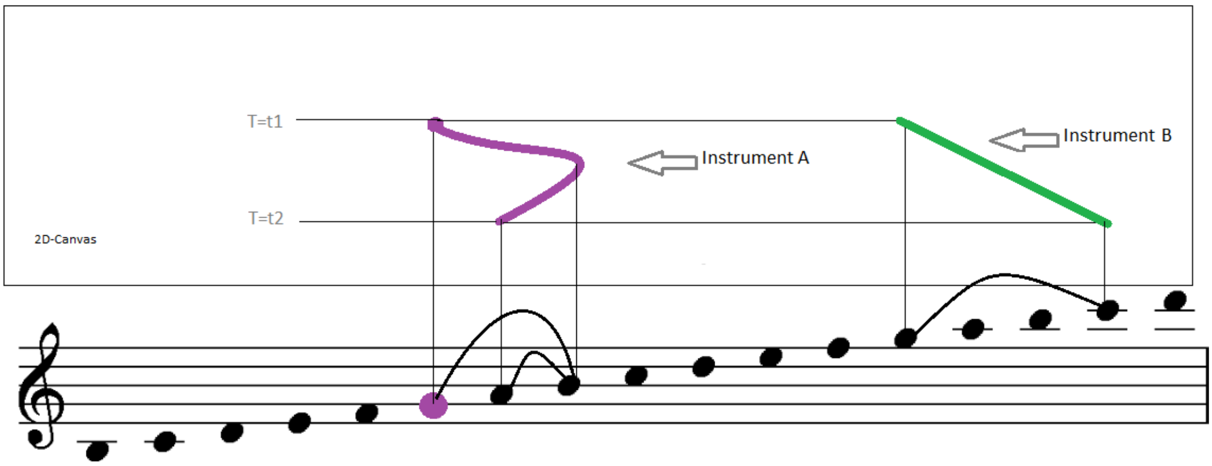


Illustration 3.3: Two instruments PFs

Lastly, the master volume is doubled in Illustration 3.4.

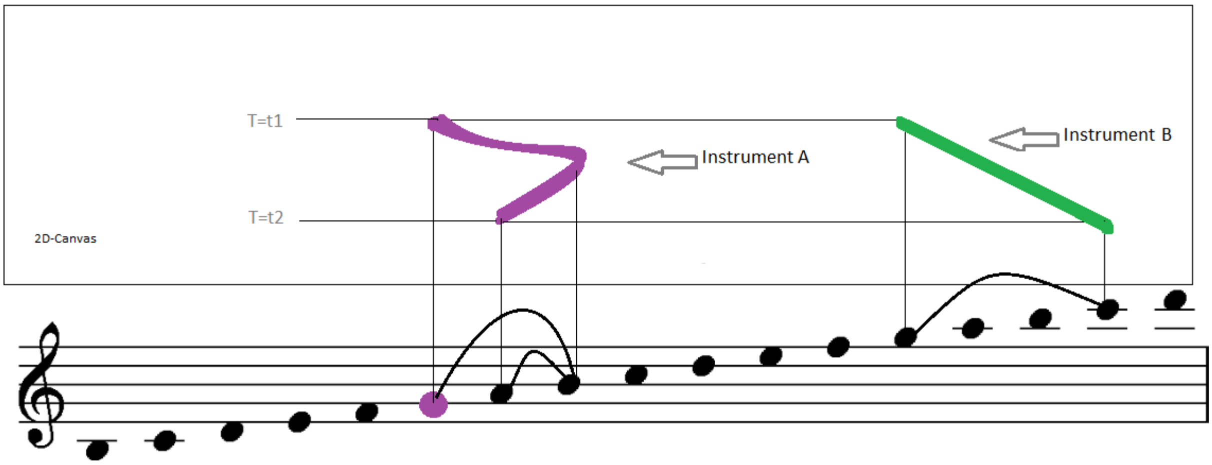


Illustration 3.4: Two instruments PFs

As shown in Illustration 3.4, each PF have had its thickness doubled. The program scans the patterns from top to bottom and records the color and thickness properties of each PF, and that is how visual patterns are converted into audio musical signals.

The other linkage shown in Figure3.1 is the relationship between finger actions and the visual features in the 2D tangible interface. This linkage transcribes into the question of how we play the visual musical game. The game design includes four modes, namely: finger drawing mode, tube mode, edit mode, replay and catching game modes. The objectives of each of the modes are as follows:

1. In finger drawing mode, users use their fingers to drawing random patterns to test the basic musical features and to give them an overview and a brief idea of how patterns are converted into audio signals.
2. Tube mode is a novel method of how computer-simulated user interface can help music creation in a tangible pattern manner. Instead of tracking direct finger movements, the Tubes are proposed in the software design, which are customized small keyboards aiming to enhance the music playing and editing efficiency. The tube mode is designed to help composers to actually create the musical melodies.

3. The edit mode includes most of the editing functions to the patterns so that the whole musical canvas can be viewed as the visual design view of the music composing. This is done by the linking finger actions to graphical manipulations, which will be further elaborated in Section 3.4 in this thesis.
4. The replay mode is a functional feature that replays the audio patterns created either by real-time finger actions or by offline editing. This mode aims to help users play the selected musical patterns or load in midi input to mix with user-created sound tracks.

With the design of these four modes, the project is able to provide a comprehensive collection of functions to create musical features and, more importantly, to achieve real-time music creation [11] which further supports multi-player and remote players, and enables saving and loading functions. The design modes shall be further discussed individually.

3.2 The Finger Drawing Mode

The layout of the tabletop user interface can be divided into three parts: the instrument selection section at the top (sections A in the figure below); the musical pattern canvas region that occupies most of the space for music representation on the pitch-time [4] conversion basis; and the tool bar at the bottom which is used to switch between modes and select functions.

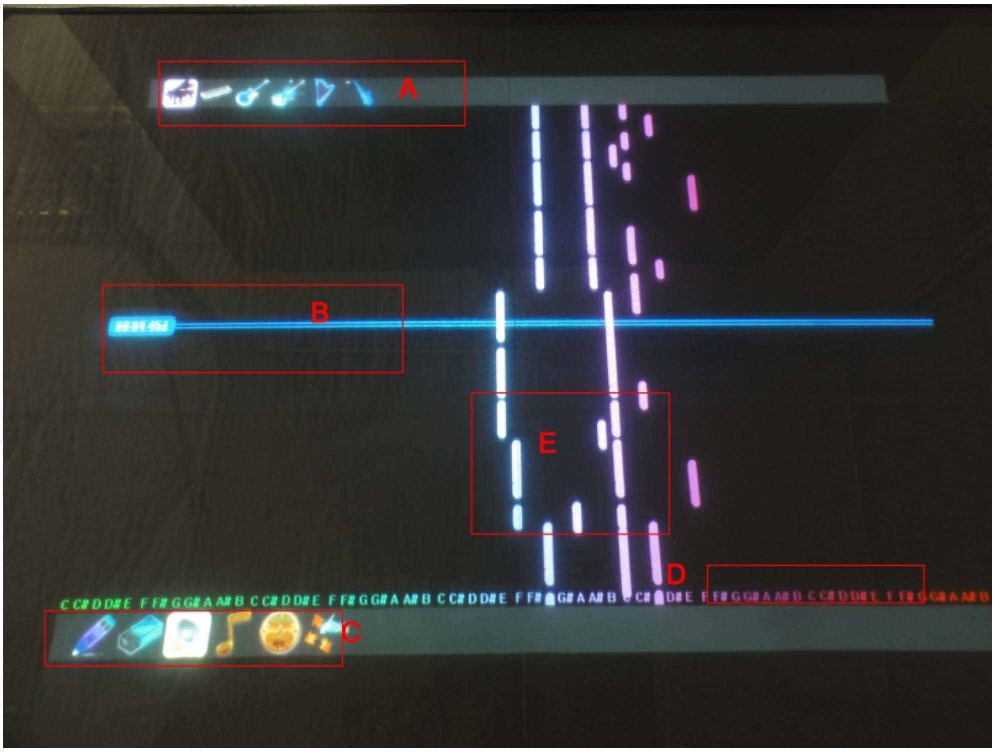


Figure3.2. GUI Design Overview

On the 2D tabletop, the top bar A is the musical tone selector corresponding to the MIDI instrument library, where users have the option to choose musical instrumental tones of PFs. The middle canvas is used to display the musical patterns such as the pattern E circled in Figure 3.2. The pitch keys (section D) are arranged in sequence with lower keys on the left and higher keys on the right, and assigned with different colors. Extended from each pitch key is a vertical grid line that is not visible in the screen snapshot in Figure 3.2. The bottom bar C is the functional bar where users are able to switch between different modes, namely: Finger mode, Tube mode, Replay mode and Edit mode, which will be discussed in the later part of the thesis. The time scanner B running from top down is the time line when the patterns are played.

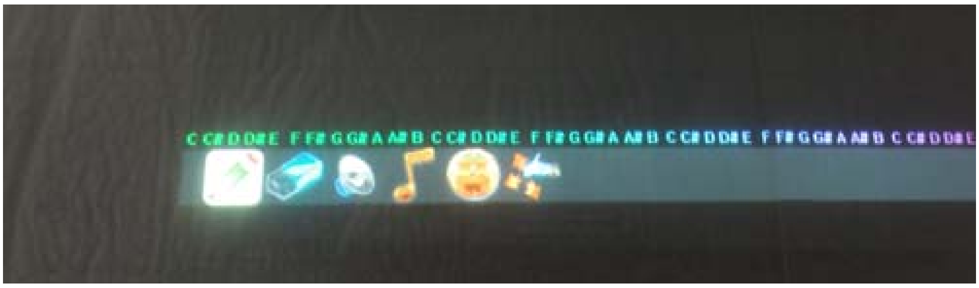


Figure3.3. Bottom menu

The finger drawing mode is the default mode for beginners to draw random patterns with fingers in order to gain a general impression of how patterns sound like. The system translates patterns into pitch and timing matrix and converts it into audio output. The output is not necessarily music depending on the how the pattern was arranged by users. However, basic music concepts can be easily inferred when the users experiment on the tabletop. Examples of these music concepts include Tone, Tone color, Compass, Pentatonic scale, twelve-tone system, and rhythms. Note that even random patterns can still be transformed into audio output simply based on pitch and timing mapping rules.

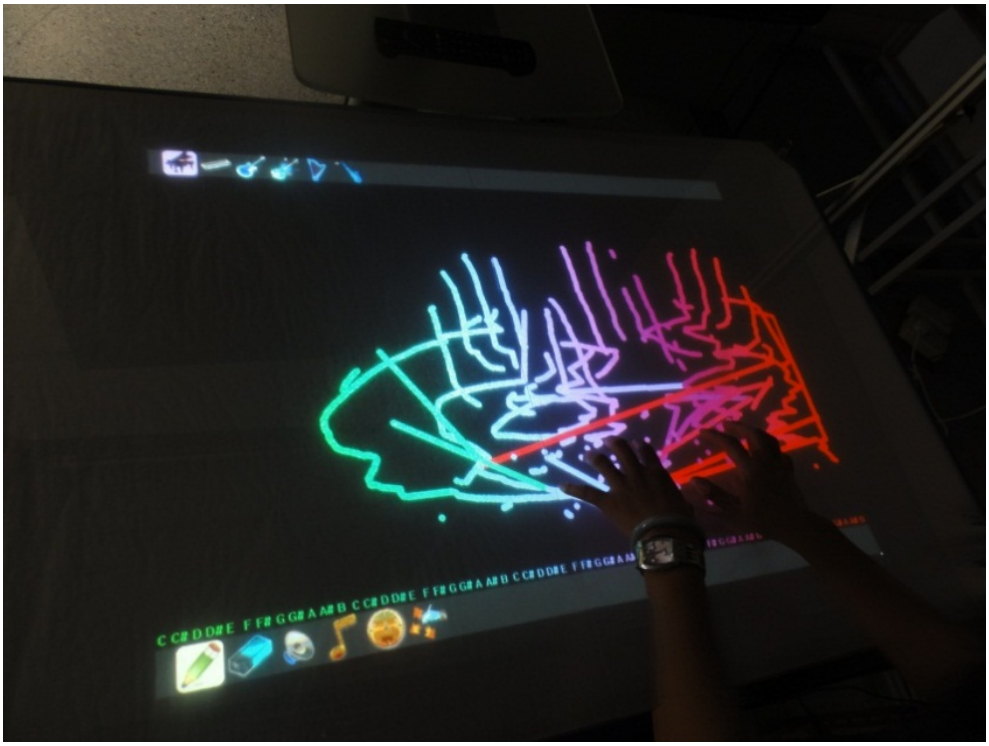


Figure3.4 Random finger drawing pattern

Under finger drawing mode, the spot of finger contact will be tracked, its location will be mapped into pitch axis, and the respective sound will be generated. However, the time related features are not so obvious from hearing and touching activities in this mode if we consider the real time sound playing (the timing feature cannot be heard but can be observed). The multi-finger touch-downs will instantly give the user the impression of pitch both from midi sound as well as the color referenced displays on the visual part.

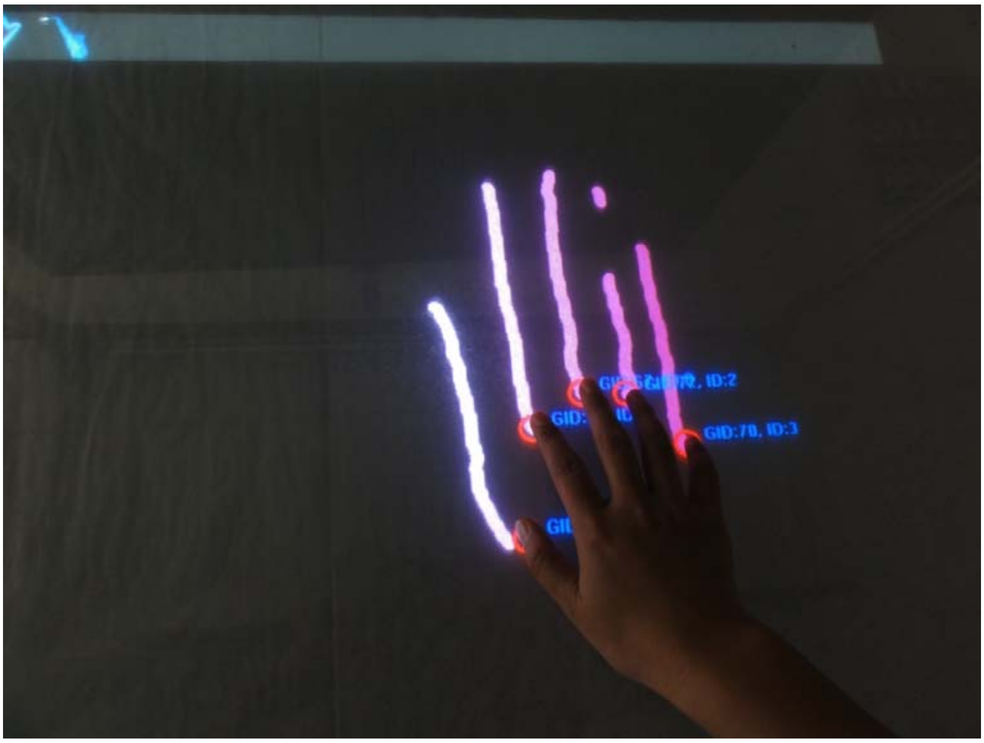


Figure3.5: finger trackings

The patterns can be as random as it can be. However, the output of a random pattern will most likely generate noise rather than music melodies. An observational study on the relationship between regular pattern and the irregular pattern and their relationship to the definition of music is described in Chapter Five of this thesis. The finger ID appearing on the above figure is for testing purpose only, and will be removed in the application execution stage.

When the patterns are played, the time scanner explicitly demonstrates the time information in the 2D patterns and how it is converted into real audio output, which may not be easily understood on the static music canvas [29]. Besides giving the novice users a chance to try the pitch time canvas layout, it can also be used for experienced user to demonstrate to the layman users the basic musical features in the time and frequency domain.

3.3 The Tube Mode

The Tube Mode is designed to simplify the finger operations in real-time music playing. The users can first select a musical tone on the top instrumental selection bar. Then, by clicking the tube button on the menu bar below, users are able to create a new small keyboard which is termed as the Tube. The Tube inherits a natural property of the musical instrument tones. The users can further test and select a pitch and drag it into the Tube. The total number of tubes is unlimited, but a proper number of Tubes should be chosen such that the Tubes are easy to operate and there still left with enough space to show the musical patterns. The next thing to determine is the total number of pitch balls within each tube. As we have five fingers on each hand, the typical range that can be covered with one hand is 6-10 pitch balls. Any tube with more than ten pitch balls will be considered difficult to operate and it is better to create another tube for operational simplicity. After all the tubes for one performer are set up ready, the preparation phase of the tube mode is completed.

In the project demonstration, the musical instruments selected are Piano, Blues guitar, Soft harp, West violin, cithara Harp and Saxophone, which are quite typical and representative. The procedure of using Tubes to create a piece of music can be demonstrated in the work flow diagram below:

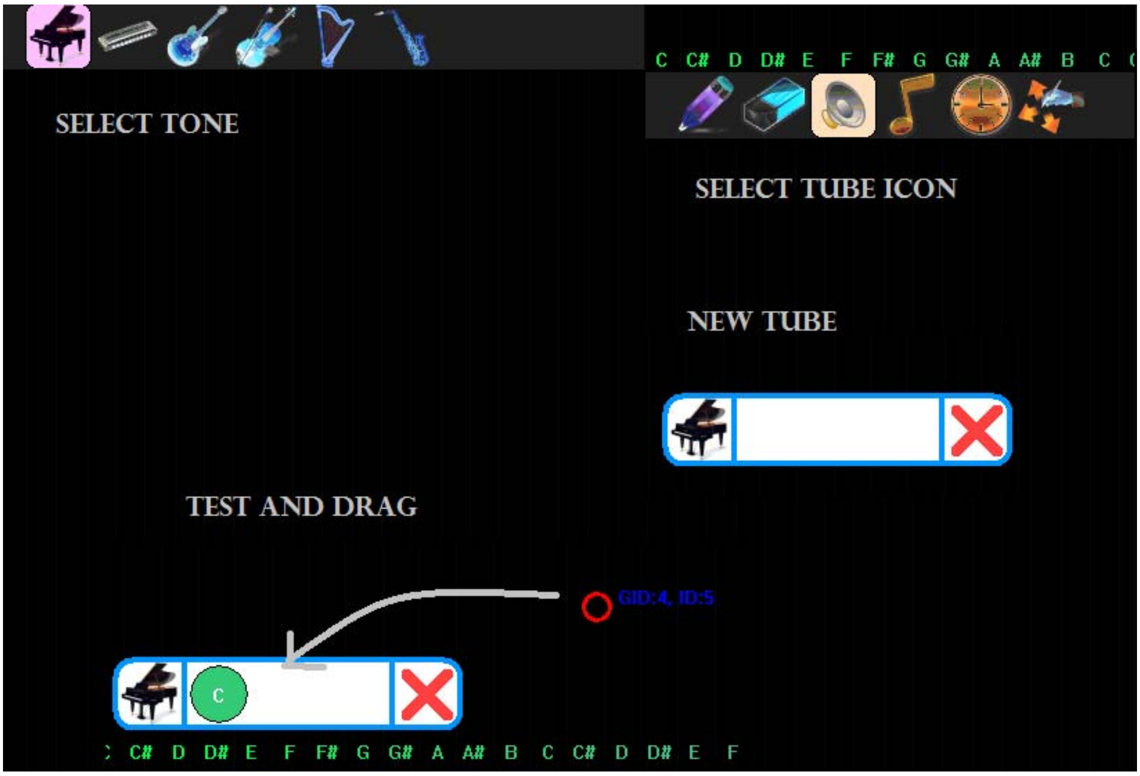


Illustration 3.5: tube mode procedure #1

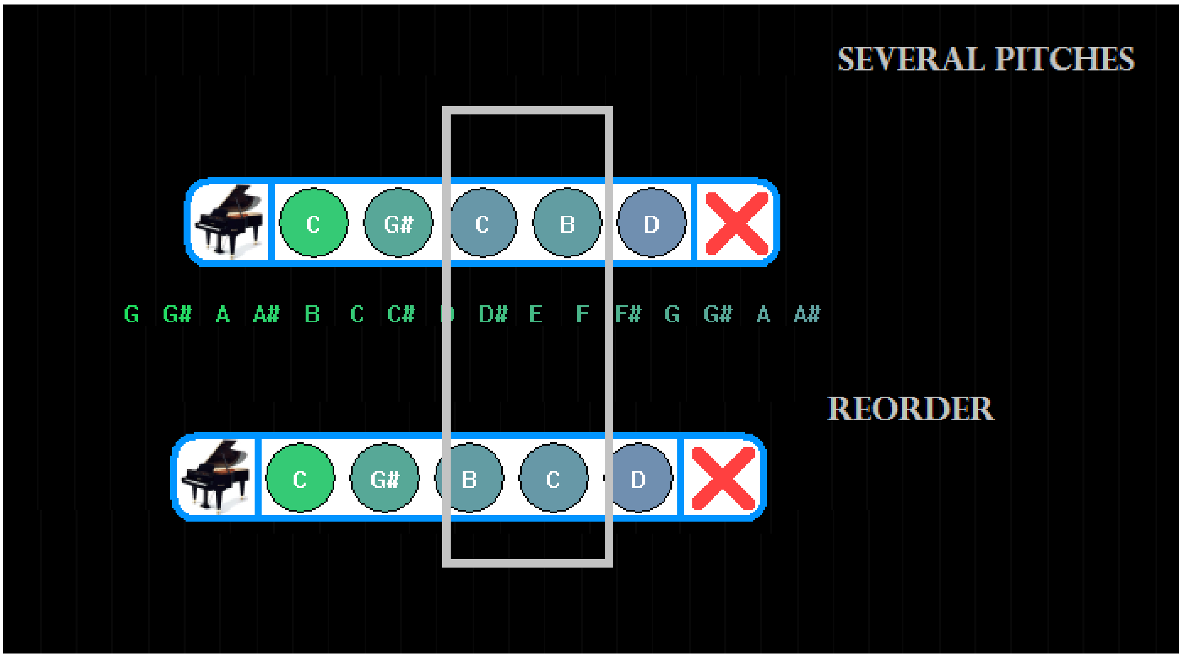


Illustration 3.6: tube mode procedure #2

Illustration 3.7: tube mode procedure #3

After the Tube group is created, it is free to be dragged to the place where the music players are more comfort with. When multiple players are participating at the same time, they need to divide the whole canvas into several regions and each region contains a certain number of tubes for their ease of operations. When the tubes and pitches are ready, there is an option for them to import other midi channel instrument to play together with the Tubes created. To proceed into the real time play, users just press down the clock item and the time scanner starts to run from top down.

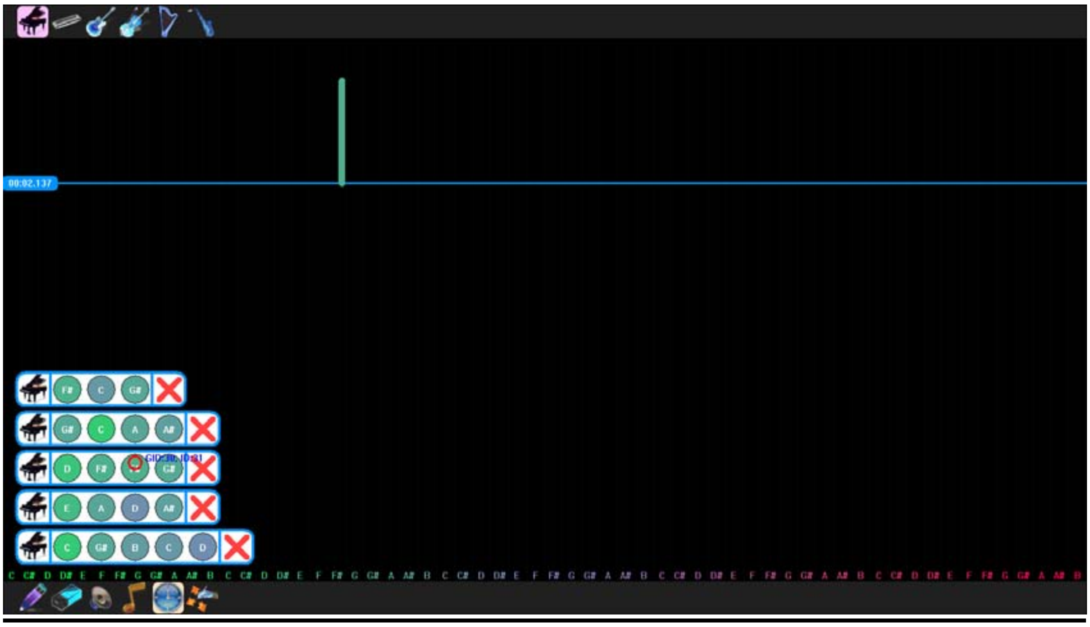
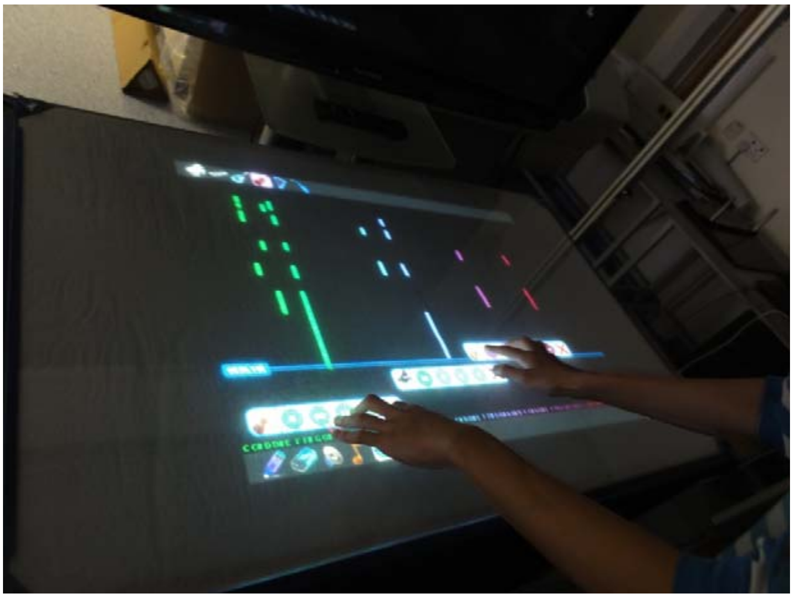


Figure3.6: time running, real time play down.

When the time scanner does not receive any input for 5 seconds, the timing mode will automatically cease. Now the pattern has been recorded and the real-time music creation has been completed. The next stage is to edit patterns which are created not exactly as desired by the composer, and to add in more features on top of the real-time recorded patterns.



3.4 The Edit Mode.

Users can enter the edit mode after creating the initial patterns on the canvas either in the finger drawing mode or in the tube drawing mode. To enter the edit mode, users can simply tab the eraser icons on the menu bar.

The functionalities under the edit mode include:

1. Redraw and overwrite the existing patterns.
2. Reshape the existing single pattern.
3. Erase the patterns.
4. Select and switch between midi channels.
5. Set the recursive playing points: start / end (for the whole channels or for a specific channel).
6. Load in music input.

   The application currently supports the input of wma and midi formats. Audio files in other formats should be converted before loading.

7. Save work as midi format.

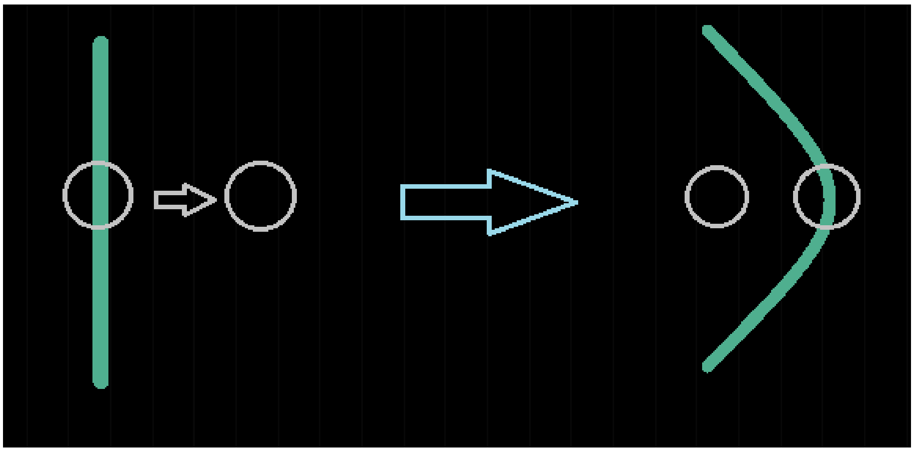


Illustration 3.8: finger operation: reshape.

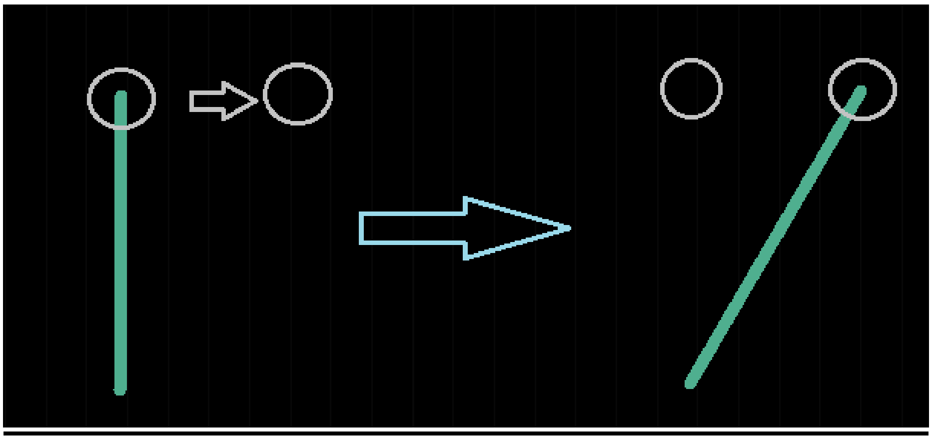


Illustration 3.9: finger operation: relocate

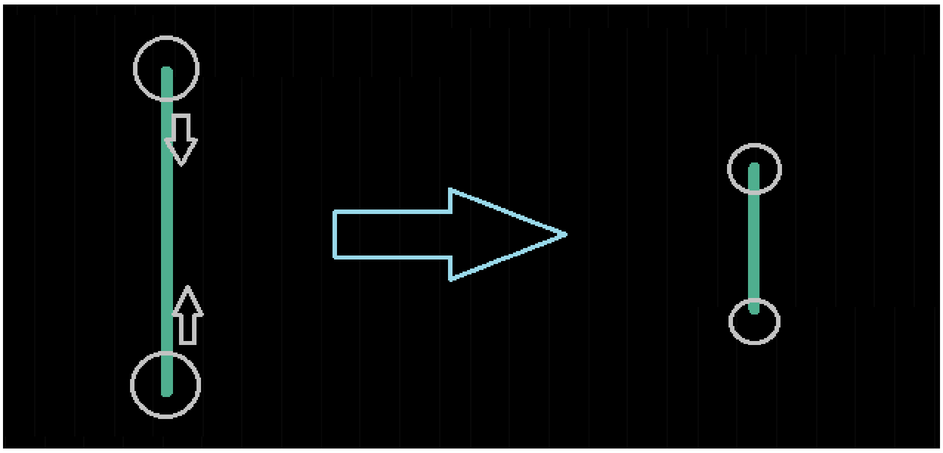


Illustration 3.10: finger operation: resize

Some finger operations, such as the resize operation, require two fingers to touch down and release simultaneously. Other finger operations, such as the pattern relocation (by one end of the line pattern) and the reshaping a lasting key to a bending key, only require one finger to operate. The tolerance range of valid finger operation areas are predefined to the ±1.2 centimeters (reference to 1024 x 768 pixels) to the core center of the patterns. In the future work, the volume control will also be achieved under this edit mode.

By pressing the “S” button on the keyboard, users are able to save the work to specified location. To load an audio file, users need to press the “L” button on the keyboard, select a directory of the file location, and then enter the numerical index for its midi channel. The index number of the midi channel for the loaded file is restricted between 14 and 30, because the default instrument occupies the first 12 channels while each musical instrument occupies another two channels: one for pitches and perpendicular pitch lasting, and the other for bending pitch achieved by pitch wheel. The drum will take five independent channels, which will be discussed in details in the implementation session.

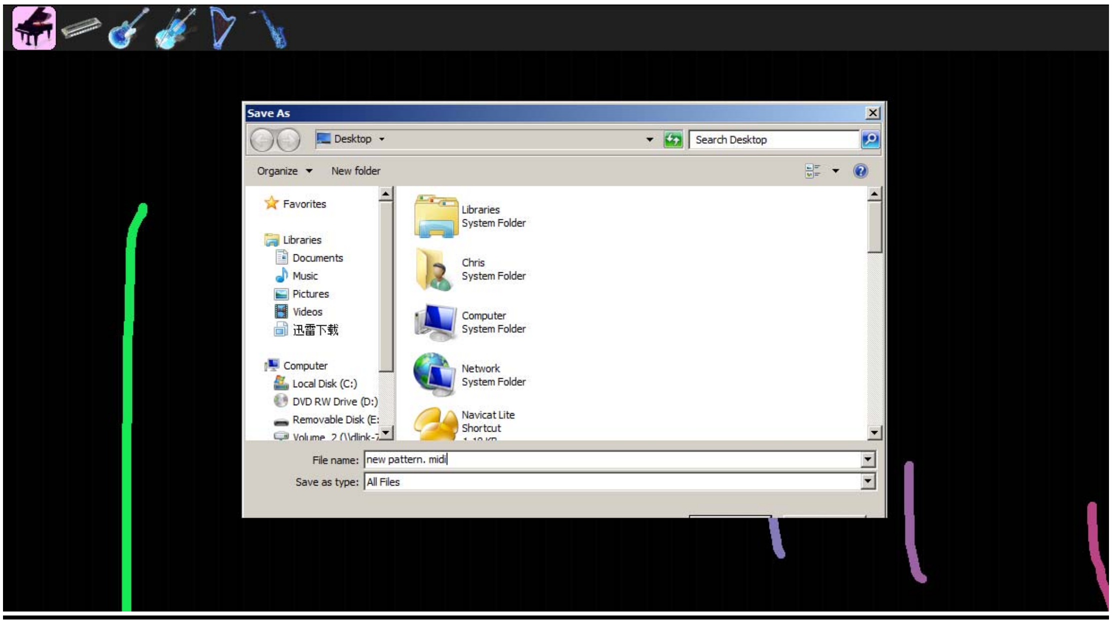


Figure3.7: save patterns

The edit mode provides most of the off-line editing tools. Comparing to the professional musical composing software, the MusicPattern software has a visual interface. However, the editing toolbox may not be comparable to professional software due to time constraint of the project. In the future design work, the toolbox is expected to include more advanced functionalities that are similar to the text-based musical composing software.

3.5 The Replay Mode

The replay mode consists of two separate sub-modes, namely the pattern replay mode and the replay-game mode. For the pattern replay mode, the system scans the patterns and interprets it as real-time midi output. Instead of the time scanner running from top down, it is fixed at the bottom in the replay mode. The canvas is rolling down, and whenever the pattern hits the time scanner, it starts to sound.

The pattern actually records the finger operation actions instead of the pattern itself. Therefore in the replay mode, all the real-time finger actions can be retrieved and replayed to the audience. For a single point pitch, the finger touch and finger release action are from the same point, and the scanner takes in the information and notes it as a point pitch. For the vertical line patterns, the scanner takes in the finger-down time spot, the respective pitch tone, followed by the ending time spot for the pattern, and then reproduces the whole lasting pitch. For the sliding and bending patterns, the scanner takes in all the instant points along the path created by the finger dragging actions. Because the starting point and the ending point are not the only parameters to define the pattern under this case, the situation is thus much more complicated for the irregular shaped patterns.

In the replay mode, since the retrieving methodologies are quite different for various kinds of patterns, it is necessary to preprocess the data and convert it into midi channel messages for the output. Strictly speaking, the reproducing procedure for the replay mode is not real-time, as the audio and the graphical animation are reproduced separately, which may lead to synchronization issue. The canvas rolling speed in terms of pixels per second must coincide with the audio playing speed in terms of beats per second (bps). Different from real-time music composing where the system simply needs to record the patterns and produce the pitch sounds solely base on pitch axis or tube balls, to playback the pattern created by users requires reading the stored four-dimensional visual features and translating it back to audio output.

The replay game mode has a different way of reviewing patterns created manually or imported from external musical files. It only replays the visual animation parts and guides the music game player(s) to hit the respective pitches on the time scanner. The pitch scales are placed right next to the scanner for user's convenience. In order to differentiate this game mode from other modes, the GUI is displayed in invert colors. When the patterns start falling from top down, the system tracks the finger operations projected on the pitch axis and generates the real-time output base only on the pitch axis finger tracking. On the visual part, only the keys falling into the valid region near the time scanner bar can be accounted as a successful finger tapping. To sum up, there is no limitation of the region of finger operation, but there is a limited region for a successful pattern finger tapping.

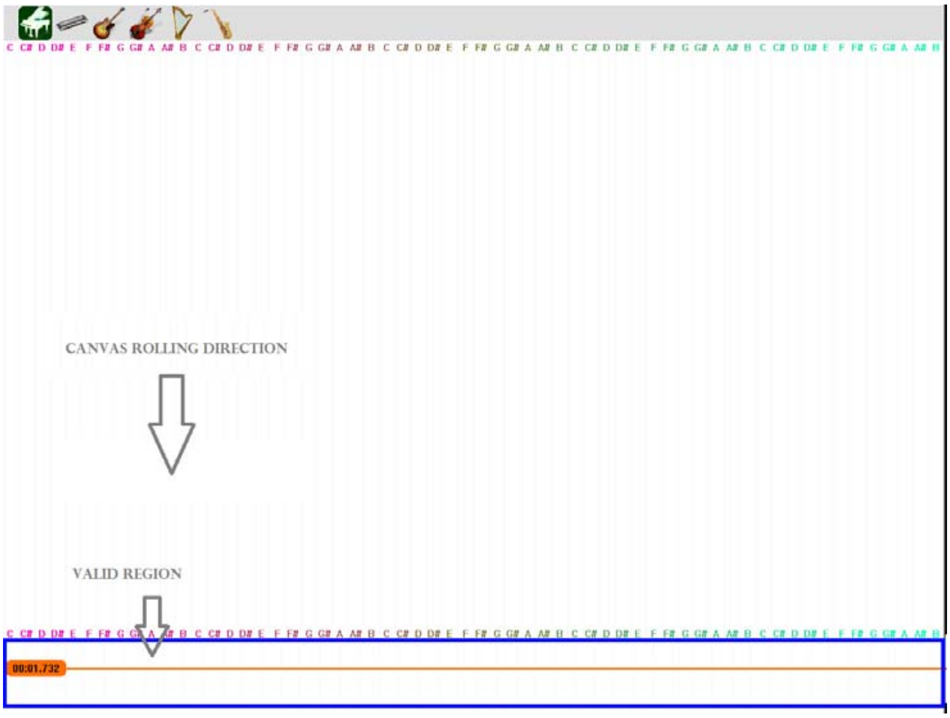


Figure3.8: replay game mode valid reaction region

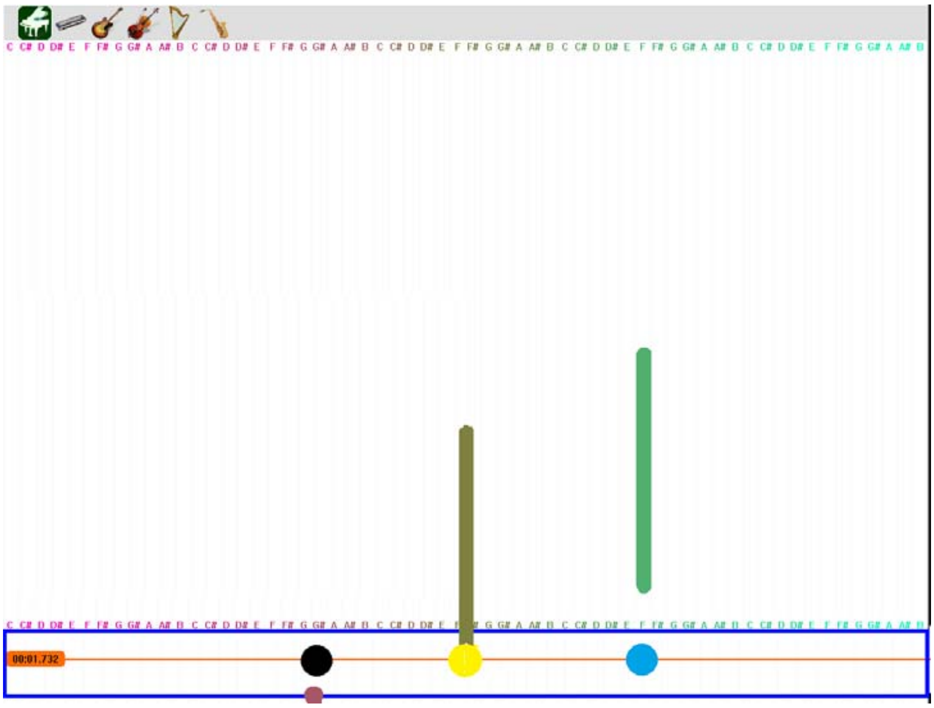


Figure3.9: replay game mode good catching

A good finger tapping is recognized when the pattern falls into the valid region and the finger touching down is detected for the corresponding pitch. A good finger release is recognized when the pattern rolls out of the valid region, and at the same time the system detects the finger released from the correct ending pitch. These will be marked as a yellow round spot.

A late catching happens when the pattern falls outside the valid region and the finger touching down of the corresponding pitch is not yet detected. A false late release happens when the pattern rolls out of the valid region, but either the finger release action is not yet detected, or the ending pitch is incorrect. False late finger release will be marked as a black round spot.

An early catching occurs when the pattern has not yet entered the valid region but the finger touching down of the corresponding pitch is detected. An early release occurs when the pattern has not completely rolls out of the valid region yet, but the finger release is detected. These will be marked as a blue round spot.

Each finger action, either down or release (dragging not included), contributes to an event, and if the event returns either a good catching or good release, it will be accounted as a successful event, whereas all the other cases will be considered as a failed event. The player's final performance will be evaluated in term of percentage based on the ratio of the successful event over the failed event. The pattern replay mode and replay game mode can be toggled by pressing the "G" key on the keyboard.

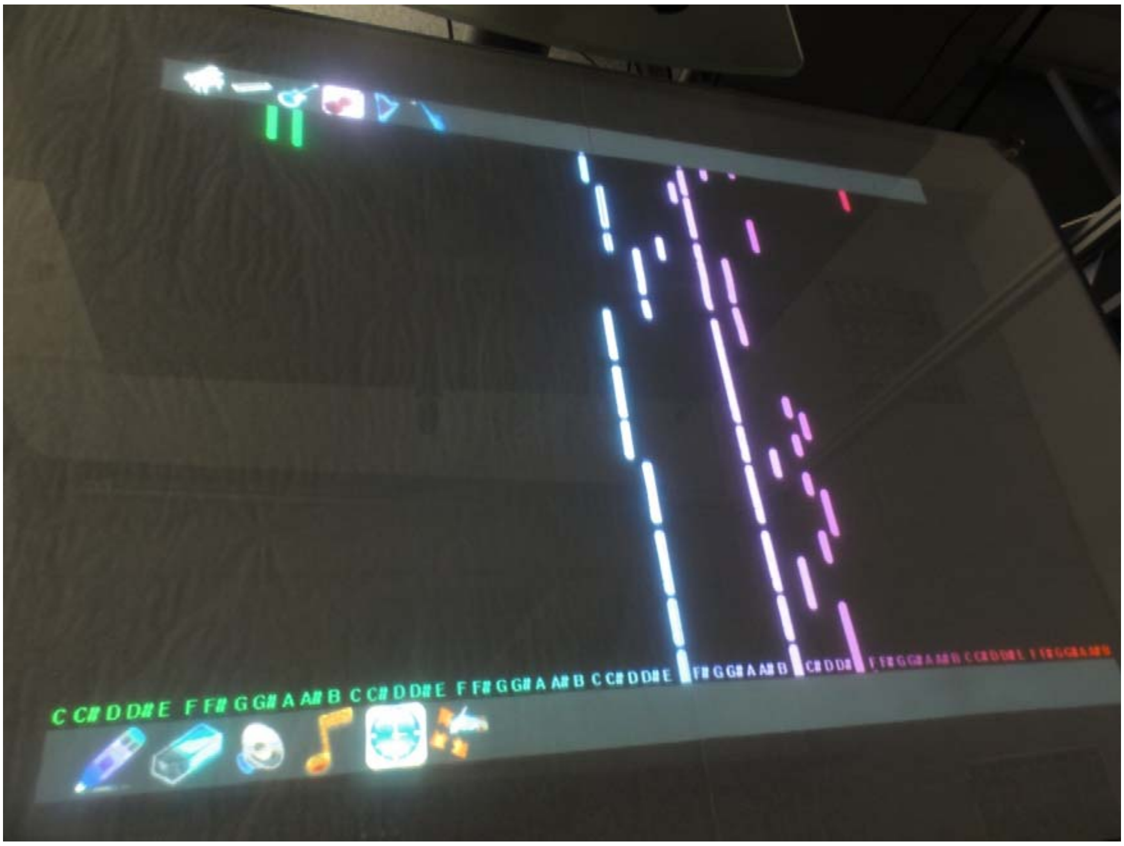


Figure3.10: replay patterns system photo

Chapter Four System Implementation

The software implementation is based on the feedback function from the TUIO camera system. The TUIO supports three main functions, which are finger down, finger dragging and finger release actions. When these actions occur, the respective function will be called in the program with the ".x" and ".y" coordinates and unique finger ID information (or multi-finger IDs). All information serves as the input data for the project's software implementation. The block diagram of the software is shown in the figure below:

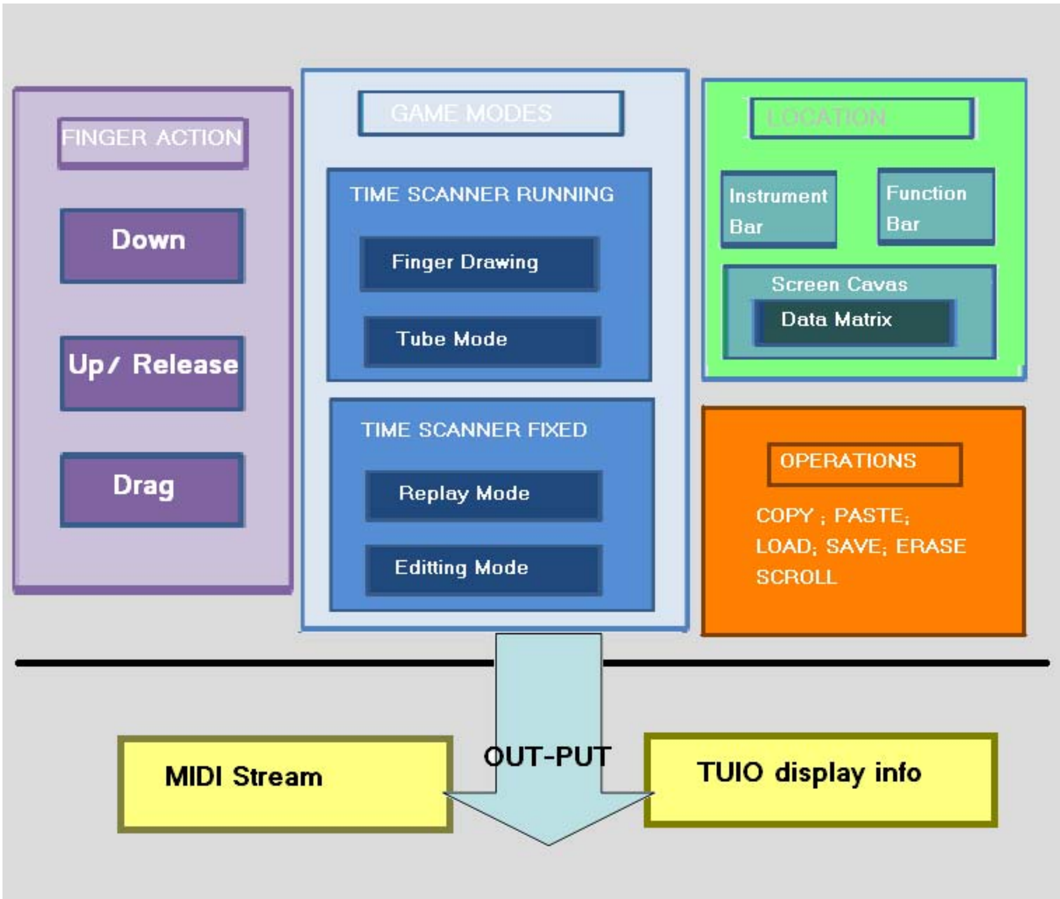


Figure 4.1: software implementation block diagram

The system output includes the visual part and the audio part. They may run under the same function to achieve synchronization or they may run under different functions and are manually set to be real-time, depending on which mode is currently on. The visual output is sent back to TUIO interface and the signal is projected back to the tabletop in real-time. The audio output will be sent through Midi channels to the computer audio card using an instrumental wavelet audio synthesizer to generate the audio output.

The project software design and logic was planned in advance of the coding. However, much more detailed issues were encountered during the coding and design cycle, among which several key implementation issues will be addressed in the rest of this chapter.

4.1 The Line Drawing Rule

In the initial design of finger drawing mode, the project does not limit the way of drawing patterns. But when it comes to replaying the patterns based on the pitch-time table conversion rule and the time scanner’s working principle, it was discovered there are a few pattern scenarios that are meaningless to the conversion for the audio outputs, such as close loop circle, horizontal lines, and intersecting lines.

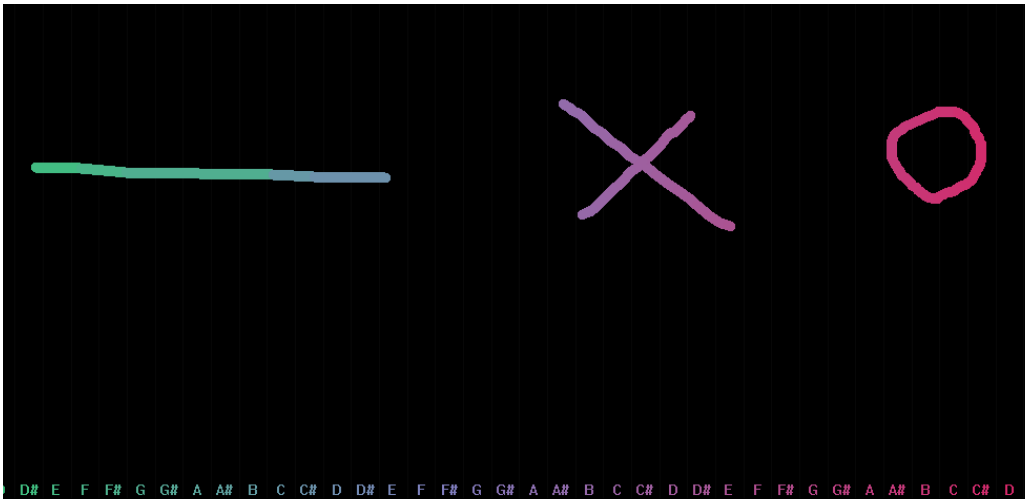


Figure4.2. Horizontal line; intersecting line, close loops

All points on horizontal line have almost the same y axis value, which translates to sounding a continuous pitch bar simultaneously. This has no physical meanings in the application, so the drawing rule defines that horizontal line cannot be drawn under the finger drawing mode.

If the same instrument is selected, it is very unlikely for one channel to sound two different bending signals as depicted in the case of intersecting lines. The midi channels are assigned such that one instrument is assigned with only one bending channel, therefore it is not possible to differentiate the two of the same musical instrument sounding at the same time. To avoid that, the intersecting lines are not allowed for finger drawing to simplify the visual display as well as to make the pattern generated suitable for the midi channel output.

For the close loops, the time scanner will hit the upper most point and then a decision has to be made whether to follow the left semicircular path or the right one. If we breakdown the loop into two semicircular paths joining at the upper most and lower most points, this will again fall into the case of intersecting lines with midi channel assignment issue. In order to avoid that, the drawing rule defines that only the downward finger movement can be recognized and stored.

4.2 Assign midi channel

There are up to 16 channels supported by MIDI. The 10^{th} channel is reserved for future work because key-based percussion is always on MIDI Channel 10, and none of the 6 chosen instruments is key-based percussion [16]. Each channel can play a variable number of sounds (polyphony). However, at any time, each channel can only play one particular instrument (sound / pitch / timbre). To utilize 15 available channels efficiently, rules are designed by scenarios.

- **Composing patterns with an indistinguishable instrument**

  The application can receive from, process and react to events from a multi-touch table. It can play, change and mute a note when players touch, move and release their fingers. When players touch the composing area on the table in PEN mode, it is defaulted to play the canvas instrument indicated on the top bar, as it is difficult to assign fingers with different instruments and track them all the time. In the TUBE mode, each tube has its own instrument which may be different from the canvas instrument. However, it is not possible to tell which tube a user will drop the note into before the finger is released. Therefore, like in the PEN mode, the canvas instrument is played when users move notes into the tubes. In both modes, a MIDI channel will be allocated to a finger contacting the touch table, unless all the channels are occupied. In this scenario, up to 15 notes can be played at the same time, but all of them are played with the same instrument. Compared with the situation of all the notes in a single channel, this approach has the advantage that each note is in a standalone channel to support effects such as pitch bending [44].

- **Composing patterns with a distinguishable instrument**

  In TIMER mode, the user can compose music patterns in real time by touching notes in tubes. The instrument of a note can be identified by the finger position and the tube instrument. The same notes playing function is shared by both this and the previous scenarios, except that in this scenario, the note is played with the tube instrument instead of the canvas instrument. The pseudo-code is shown as below.

```
FUNCTION playNote(noteOff, noteOn, pattern, noteInstrument)

  IF MIDI device is not ready THEN
    Return
  ENEFOR

  IF finger down THEN
    FOR each channel of 16 MIDI channel
      IF the channel is 10th channel or has been assigned to a pattern THEN
        Continue
                                          ENDIF

      Assign current channel to the pattern
      Play current note
      Break
    ENDFOR

  ELSEIF finger move THEN
    FOR each channel of 16 MIDI channel
      IF the channel is 10th channel or not assigned to this pattern THEN
        Continue
      ENDIF
```

Figure4.3. Code solution to the instrument confusion problem

- **Playing patterns**

  In PLAY mode, it is not necessary to play the notes in a pattern by triggering the note on and note off messages because no touch event is required from user [38]. In order to play patterns smoothly, all the patterns are compiled into a MIDI stream which is played with the MIDI stream API. Because the pitch bending applies to a channel rather than a note, any pattern that has a pitch bending should be in a separate channel by itself. When compiling patterns to stream, the first 6 channels are allocated to the 6 instruments on the top bar of the application, and the remaining 9 channels are dynamically allocated to pitch bending patterns, while the 10th channel is still reserved.

4.3 Designing Scrollbar

At the beginning, the top and bottom bars were used to scroll the canvas. It is straightforward for users to scroll the canvas by sliding their fingers in any empty area of the top or bottom bars. However, it is not easy for some users to physically reach to the top bar.

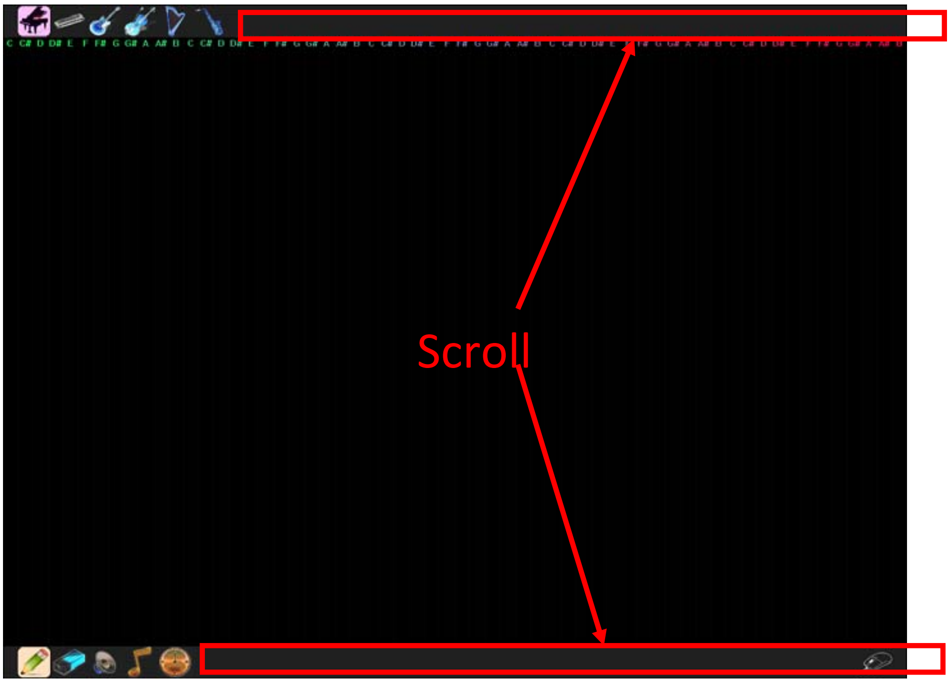


Figure4.4. Scroll bar sides

The second version came with a scrollbar at bottom bar where users could slide their fingers to the left or right in order to move canvas up or down. The idea was borrowed from some Apple products like iPod Nano whereby users can move their fingers clockwise or counterclockwise on the outer circle of the control buttons to scroll through the menu items, adjust the volume or move to any point within a song [19]. Since most of the commercial applications place their vertical scrollbar on the right border, users may feel counterintuitive to move canvas vertically by dragging a horizontal scrollbar. The reason for not installing a scrollbar on right border is to save up more space to the composing area, because otherwise the GUI (Graphic User Interface) would appear compressed.

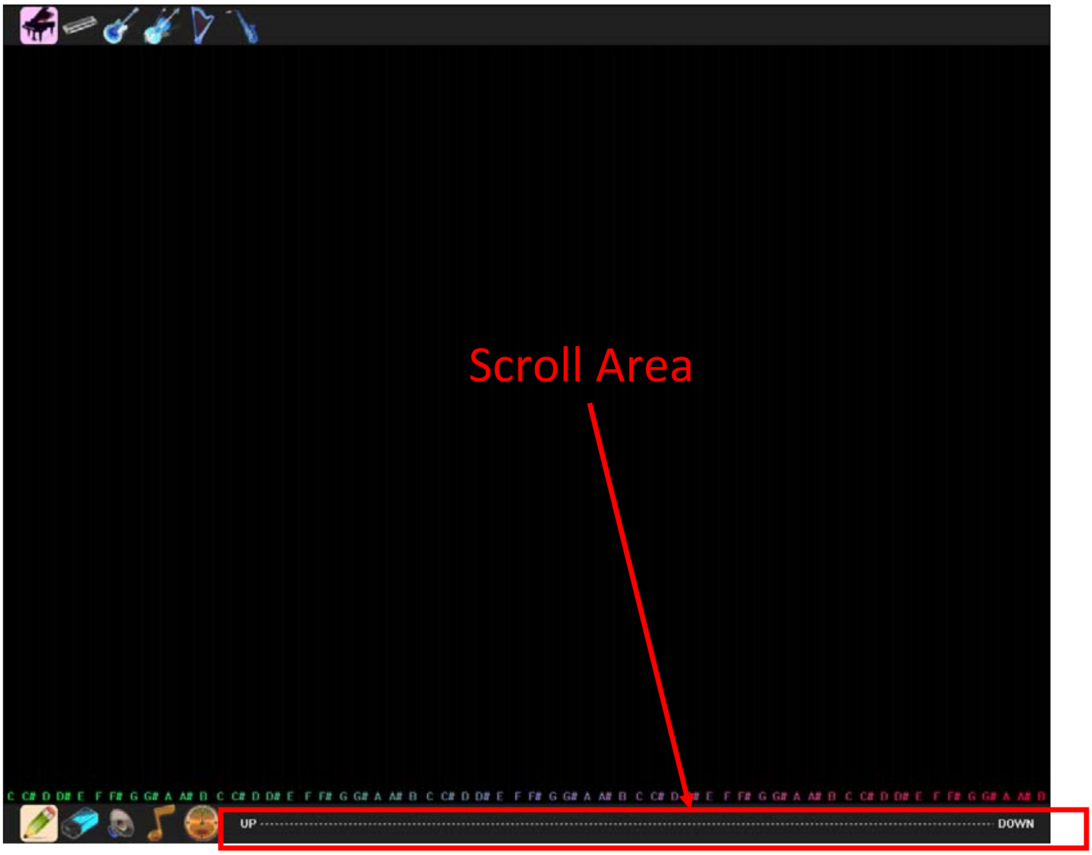


Figure4.5. Scroll bar bottom

In the latest version, a SCROLL mode is introduced in the bottom bar. In SCROLL mode, user can drag the canvas directly with their fingers. If a user drags the canvas with one finger, the canvas moves following exactly how the finger moves. When multiple fingers drag the canvas towards same direction, the scrolling effect will also increase multiply.

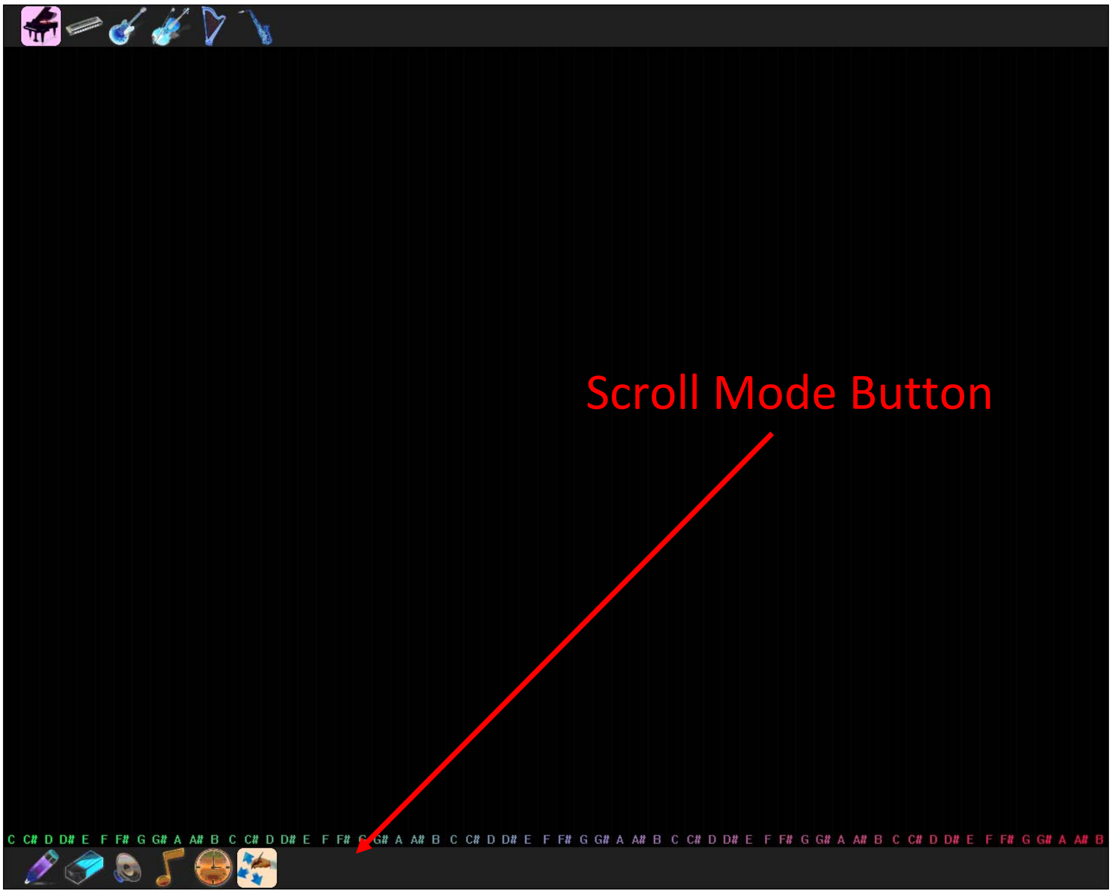


Figure 4.6. Scroll button

4.4 Designing Tubes for Composing Music Patterns

The tubes are designed to hold some notes and are used to compose music patterns. The motivation behind this is that a song is usually composed with a subset of the notes. It is possible that a beautiful melody employs only a few notes. Therefore, tubes with limited notes are more efficient than a canvas with full five octaves.

In the first version, tubes were defined by drawing a single pattern on the canvas. After a tube was filled, it could be used in TIMER mode. With the timeline showing on the canvas, whenever a note was clicked, a key point was created for this pattern which was constructed by connecting all the key points. During the composition, any notes in the tube could be clicked, and the used notes would be removed.

Figure4.7. Stocking tube

Later, it was found that users usually put the desired notes into the tube in order and seldom made a pitch bending pattern. The behavior of the tube was modified accordingly: only the first note in the tube could be pressed to start a pattern. Once the first note was released, it would be removed and the pattern was ended. To make a pitch bending, users drag the first note up or down during the composing in TIMER mode.

Figure4.8. Tube timer

In the latest version, the tube was improved according to user feedback. A tube consists of three parts. The left part shows the tube instrument icon, and by dragging it the tube can be moved on the canvas. The middle part shows all the notes contained in the tube. The right part is the button to remove the tube from the canvas. Besides the GUI, the behavior of the tube is also improved. It no longer supports pitch bending by dragging the note up and down, because unwanted pitch bending patterns might occur if the finger wobbles. All the notes can be clicked to compose patterns, and the used notes in tubes will not be removed.

Figure4.9. Tube Small keyboard

Chapter Five User study

5.1 User Groups

In our early user study, we invited 5 students from Nanyang Academy of Fine Arts (NAFA), School of Music Composition Studies, and 6 laymen users from the engineering department in local university. The overall participants' reaction was very positive. The idea of using visual patterns to represent musical features is well accepted. The interface was perceived as very intuitive since the concept of pitch keys and timing does not require much learning to understand. Among the participants, the five NAFA students are from the School of Music Composition Studies, and they have an average 8 years of music composing experience and are very familiar with the music composing related terminologies and techniques. They are treated as the experienced group of users for the MusicPartterns tests. The other 6 testers are from National University of Singapore, School of Electrical and Computing. They listen to music as a hobby but barely have any understanding on music related concepts or music composing experience. They are treated as the layman group users for the MusicPartterns tests. These two groups will be assigned with three categories of tests and the observations and result analysis will be discussed in the rest of this chapter.

5.2 User study One – the " Organized pattern"

5.2.1 Design Objective

The objective of this experiment is to study how the patterns, organized with straight lines, dots, and curves, can be appreciated by both layman and musically trained users. To the layman group users, the experiment is design as an educational function where users are able to appreciate the musical features with visual aids on the pitch-time canvas. For the experienced NAFA student group, on the other hand, observations on how patterns can help them study the definition of music, chords, and voices from the tangible visual patterns will be stated in the following paragraphs.

5.2.2 User study #1 for Layman group

Experiment

The pre-experimental questionnaire was given to the layman to screen out those who have known some of the musical concepts to be used in this experiment. We left with six layman users who may have heard about some of the music concepts but do not actually know them, such as chords, pitch bending, pitch sliding, etc.

In this user study, the layman users were first given a short introduction on the pitches and how it can be represented on the application tabletop. The pitch layout is exactly the same as it is for the piano, i.e. from low to high containing seven octaves. They tried to draw finger patterns under the finger drawing mode to test how patterns would sound. This gave the layman users a first impression of the music pitch-time table.

Next, the sample musical features were displayed on the tabletop, and they were played for the layman users to hear, while they can also see the time scanner running through those patterns. The patterns with simple features were shown first, and moving gradually to more complex features. After the demonstration ended, the new patterns were presented to the participants, and they were asked to reproduce that musical feature with their own voice. The participants were allowed to touch down the tabletop to hear the point pitches but not line patterns. The test patterns and the results are shown in the table below:

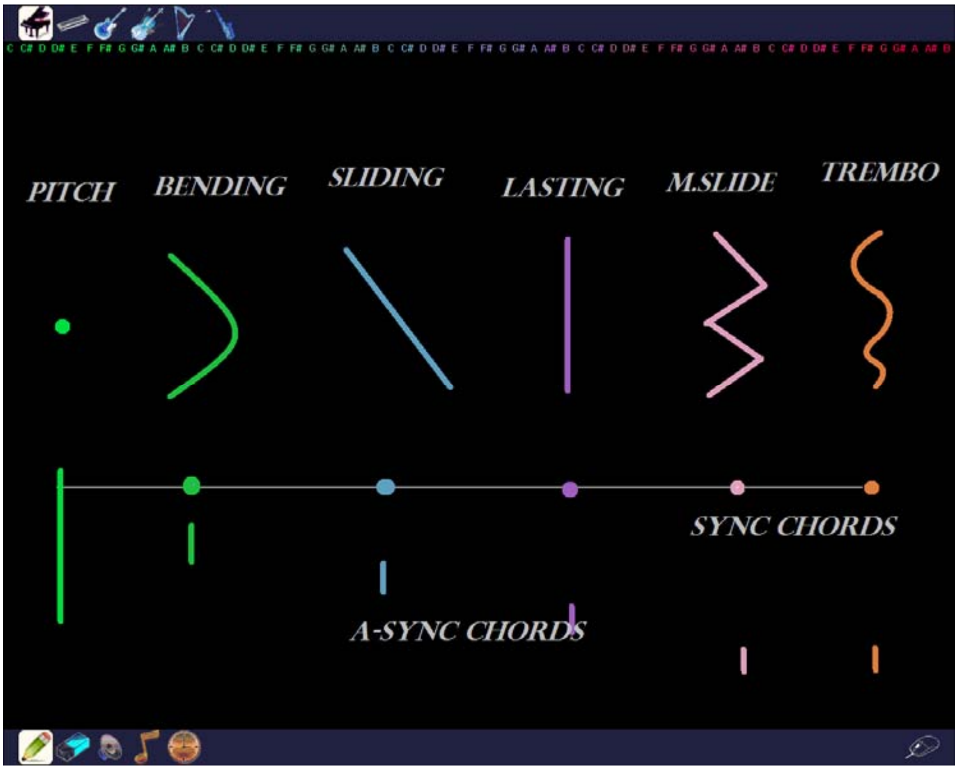


Figure5.1. Common Features Representation

<u>Result and Analysis</u>

Users	Point Note	Lasting Note	Sliding Note	Bending Note	Vib Note	Sync Chords	A-sync Chords
User A	F.C	F.C	F.C	P.C	P.C	Fail	P.C
User B	F.C	F.C	P.C	P.C	Fail	Fail	Fail
User C	F.C	F.C	F.C	F.C	F.C	P.C	P.C
User D	F.C	F.C	F.C	P.C	P.C	Fail	Fail
User E	F.C	F.C	F.C	F.C	F.C	P.C	P.C

LEGEND
F.C: Fully Correct; P.C: Partially Correct; Fail: Fail to replicate.

Table 5.1: the experimental #2 result for laymen group

Given 3 simple patterns which are pitch bending (left half of circle curve), pitch lasting (straight vertical line) and a 4-pitch G chords (4 points laying on the same horizontal line), participants were asked to guess what the patterns sound like. This task is difficult for some users but easier for others. However, all of them were able to get the first two patterns correct; most of them understood the patterns of the sliding and bending pitches, which are represented by the diagonal lines and the curves. For the chords, it is impossible to voice by single person, and 4 out of 5 participants realized that fact. The instructor further asked the participants to voice the chords by singing together. They were able to control their own pitch at the beginning, but eventually getting interfered with other participants' voices and singing the wrong pitch for the chords.

At the end of this experiment, the patterns were given names to the participants and they claimed to understand the musical concepts via this learning process supported by visual patterns.

This experiment shows a phenomenon that for some layman users, they are able to appreciate the music concepts without even knowing the terms, which matches our original project design goal: to provide an intuitive platform for layman users to play and appreciate music.

5.2.3 User study #1 for musically trained group

Experiment

Edgar Varese once gave a definition of the music as "organized sound", and this seems like a good place to start, more because of what information such a definition does not give than for what it does give. The definition is somewhat satisfying because it seems to subsume practically anything that is called music that anyone could conceive as music and certainly anything that might be called music in the future. On the other hand it is unsatisfying because there are no clues in the definition on how music is constructed. When Mulling over Varese's definition, the first question that comes to mind is an abstruse one: what is organization, and answer the question of how the sound is organized?

The more advanced music features are the melody features, where the patterns combine both the timing beats information as well as the pitch flow information. What type of pitch combination will essentially form up the voicing music and what combination is essentially not, is a rather a hot discussion area in the music composing academia. However, there is no clear definition of what is the musical motion features from the pattern recognition view because each individual has a different perspective of how audio melodies can match their own emotion feelings and appreciation. On the other hand, there is no clear differentiation between music and noise, although it is to the majority group of people that randomly created patterns are not considered as music.

Result and Analysis

The key of this test is also to test the effectiveness of the visual aids of the MusicPatterns, which has constructed a pitch-time table-format layout. The target of this test is to show the effectiveness of the tangible interface to the observation of musical concepts, with focusing on the advantage to display the musical melody features.

The layout of major Chords, namely Cmaj, Dm, Em, F, G, Am, G9 are represented in the following figure:

Figure5.2. Seven frequently used C class chords

The major chords in C-maj, Dm, and G, Am follow the same pitch interval. But the general trend is that the base pitches are shifting one pitch to the next pitch in the octave model as they define the chords. The second pitch alters between One to Three intervals which will result in different emotions. Among them, the second scale representing the most “complete and comfort” pitch which will be accounted into the major chords. The third pitch in the chords is in the high pitch frequency within the respective chords. We are more sensitive to the high pitch region in the chords, and the pitch lies between 4^{th} Scales to the 7^{th} Scale. The fourth pitch tone is the next octave repeat from the base pitch tones. The basic formation of simple chords only consists of three pitches. The fourth and higher pitch is optional and not decisive in the pitch emotion perceptions.

There are visually standard formats of architecture, such as sonata allegro form, fugue, rondo, etc., all with a set of large scale rules. At the other end of the scale just above the level of pitch, is the level of timing. This is where we can talk and work comfortably because the complexity is manageable. Moreover, to be able to control larger structures, one must be able to control the smaller structures from which the larger ones are built.

The tempo beating controls can be viewed from the following figure:

Figure5.3 Rhythm notations on canvas

The common musical rhythm can be used in musical classification. If we only consider the tempo information from the table canvas, the detailed tempo pattern can be shown in the above figure. The time scanner will be running from top down to reproduce these tempos, the pitch locations is not important in testing tempo rhythm in this case. The MusicPattern under such test has well represented the patterns that can distinguish the different musical class types as in demonstration. The patterns are clocked by finger touch down and releases. To use finger of timing control is also one of the advantages of the MusicPattern which will also be discussed in this chapter.

5.3 User study two- The Timing Control Test

5.3.1 Design Objective

The Objective of this test is to study how efficient the finger tapping control of the rhythm and tempo related musical features over text editing. The text-based notations include the chords, pitches, and a set of musical instrumental specialized text notations such as guitar's six spectra and drum notations. It is a common fact when we are trying to re-synthesis the music melodies that the most difficult issue to handle is the control of the melody rhythm. However, it is something that we can easily clock simply by finger tapping actions for even less musically trained users.

5.3.2 For Layman group

<u>The experiment</u>

The experiment first produced 5 different styles of musical rhythm, and then the layman users was asked to use finger tapping to clock on the tabletop surface to reproduce the respective rhythm, and then use the replay mode of the MusicPattern to check the correctness of their reproduction. The result can also be visually checked from the canvas too. A tempo matching code is used to return the percentage of the correctness of the finger clocking by the user.

The finger tapping recognition code first takes in the sample and save its y axis (timing axis) information as well as the length of the pattern. The end of pattern is auto detected by the silence recognition for a predefined time interval. The program then takes in multiple users' input and stores them in respective arrays. In order to compare the percentage of each user's tapping correctness, the system compares the finger tapping record against the sample rhythm characteristic. The code only reads the timing information of the points for finger touch downs and finger releases. Whenever these two actions are not synchronized with the sample pattern, there will be an error contributing to the overall performance, and this is recognized by multi-scale processing error [40]. The program will return a percentage from 0 to 100.

Result and Analysis

The test results of the first rhythm (jazz style) are shown below:

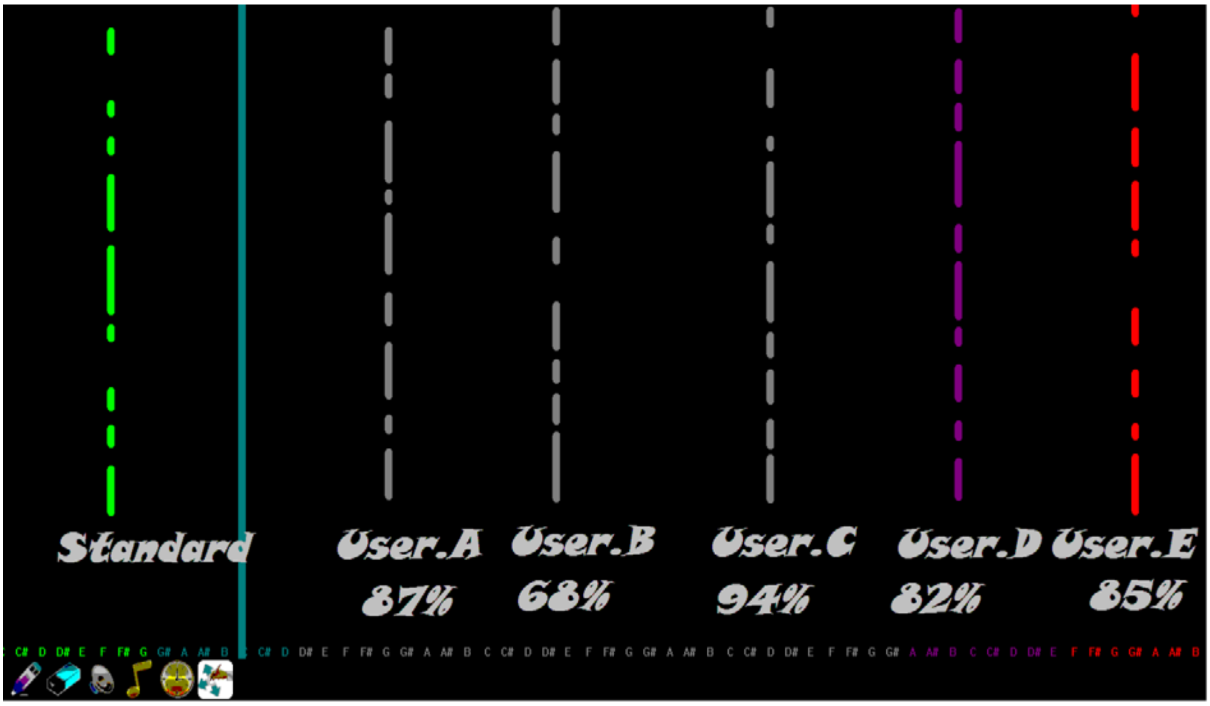


Figure5.4. Layman rhythm tapping result demo

There are five time clock tests, namely Jazz, Waltz, Classic pop (one of the representative), Samba and Rock style. The overall result can be shown in the table below:

Scores/ Participants	Jazz.T	Waltz.T	C.Pop.T	Rock.T	Samba.T
T1	87	89	92	91	84
T2	68	77	88	75	81
T3	94	93	95	89	95
T4	82	90	88	83	91
T5	85	85	92	94	96
Overall	83.2	86.8	91.0	86.4	89.4

Table5.2: First time replay.

Scores/ Participants	Jazz.T	Waltz.T	C.Pop.T	Rock.T	Samba.T
T1	97	98	98	96	93
T2	89	98	94	92	91
T3	91	92	88	91	93
T4	96	100	98	94	90
T5	93	94	100	99	92

Overall	93.2	96.4	91.0	94.4	91.8

Table5.3: 4th time replay.

There is a clear observation that after the participants hear the sample rhythm over and over again they can memorize the rhythm better and better. There are several reasons for this. The most important reason is that they get more and more familiarized with the sample rhythm and can remember much better than the first time, and they are also getting more familiarized with the finger clocking practice. As a matter of fact, for the musically trained users, they are able to achieve 90 percent on their first shot easily. However, it is also difficult to achieve a hundred percent correct. This is due to the hardware catch up limitation, and the requirements for an exact correct case are very stringent and are not that important in this user's test.

For layman users, this test demonstrates the role that the MusicPattern can play in educational activities for tempo-related musical concepts. The fact that they are able to reproduce certain pattern of different rhythm styles shows that they already learned the key rhythms of that style. The experiment also shows that even layman users are able to achieve a high percentage of correctness of tempo performance after repetition of hearings and trying. This test result will be compared to the first composing tempo with text-based notations for musically trained and experienced groups [3].

5.3.3 For experienced group

Result and Analysis

The main objective of this test is to show the ease of using the tempo information control compared to composing using text-based notations for the musically trained composers. The musically trained composers have much better hearing and memorizing capabilities than those of layman users. They are able to memorize any new tempo rhythms within one or two times of listening. However, the memory may not last for too long. The information stored in their mind is not converted into the notations for recording until they write it down, and when they are trying to mark it on the text-based stuff [13], they always get confused by the written notations from professional text-based software: the reproduced audio rhythm by the software is not quite as what they think it should be. During the modification process, they may experience loss of information for the rhythm that they are supposed to record or create in composing.

In contrast, when they hear the tempo rhythm for first time, they can easily tap it on the tabletop. Tracking their finger tapping and converting it into text-based notations is one of the functions that the musical related HCI applications can do. The MusicPattern integrates the direct finger operations and will be tested to show its capability in helping the composers to master the tempo flow.

The first task is to compare the accuracy of musically trained composers who are using text-based professional software to record the tempo flow with the layman user's finger clocking results. To make this comparison fair to the layman groups, participants are first given to five alternated random rhythm signals that they are not so familiar with. Because the standard rhythm styles are already very familiar to the musically trained composers, testing them with the meaningless tempo rhythms is to make sure that the level of strangeness is equal to both groups.

The comparative result is show in the figure below:

Scores/ Participants	Jazz.T	Waltz.T	C.Pop.T	Rock.T	Samba.T
P1	60	90	80	90	70
P2	70	70	70	100	60
P3	75	80	90	80	70
Overall	68.3	80	80	90	63.3
Layman's Result:	83.2	86.8	91.0	86.4	89.4
Difference	-14.9	-6.8	-11.0	3.6	-26.1

Table5.4 the first time of hearing on tempo control

Scores/ Participants	Jazz.T	Waltz.T	C.Pop.T	Rock.T	Samba.T
P1	70	90	70	80	80
P2	70	70	80	100	70
P3	65	80	90	80	80
Overall	68.3	80	80	86.6	76.6
Layman's Result:	93.2	96.4	91.0	94.4	91.8
Difference	-24.9	-16.4	-11.0	-7.8	-15.2

Table5.5: the fourth time of hearing on tempo control

It can be clearly seen that even the musically trained composers achieved less accuracy in tempo information recording than the non-composers! Although it is certainly not fair for the composers to use text-based notations, this comparison shows that the finger tapping is more useful for tempo recording than text notation with equal amounts of rhythm information.

Another practical test was given to the other three musically trained students where they are ask to compose the given tempo rhythms and the time needed for them to achieve the hundred percent correctness are recorded.

participants	Base Beat	Sole tempo	Chords Specs
P4	12s	1min 20s	5mins
P5	9s	1min 50s	6mins
P6	17s	2 min 20s	9mins

Table5.6: The GuitarPRO software (Text-based) (Finger)

participants	Base Beat	Sole tempo	Chords Specs
P4	53s	3min 10s	4mins
P5	1min 10s	2min 15s	7mins
P6	40s	2 min 40s	8mins

Table5.7: The MusicPattern (Finger)

This experiment also shows that direct multi-finger actions has better efficiency in recording the tempo related musical concepts. To support finger operations in music creation is to further facilitate the composing process. With this design idea in mind, the interactive tabletop design is able to help the composers create and re-synthesize music.

5.4 User study Three- The music melodies.

5.4.1 Design Objective

To fully enjoy the music creation, the end users must be able to create real music rather than merely create or understand the musical components. In this experiment, layman and musically trained participants were assigned with the task of creating music but under different composing mode: novice participant were asked to play the music in the replay game mode where only the tempo component are controlled. The experienced users were asked to create and compose real music melody segments. Survey was given to both groups for their feedbacks.

5.4.2 For Layman group

The Experiment and Discussions

Similar to Guitar Hero games, under replay game mode, novice users were asked to catch the falling patterns when it hit the bottom. Different from the other musical time catching games, the MusicPattern game mode is unique in the following aspects: a) real pitch axis. The pointes on x-axis have a one-to-one mapping that of a piano; b) the game platform supports multi-players and multi-fingers. This will contribute to the enjoyment of group collaboration and communications; c) the music patterns created with the MusicPattern application can be played in the replay game mode. This means the creator of the music patterns could ask his or her friends to enjoy the music he or she created by playing this game.

Novice users in this experiment played the game melodies that were created by the experienced group. The correctness of the catching game play is shown in the table below:

Scores/ Participants	**Melody_1**	**Melody_2**	**Melody_3**
P1	92	79	84
P2	87	86	94
P3	89	93	97
P3	78	81	88
P5	91	96	93

Table5.8: 37 seconds music replay game results

As this game is designed purely for entertainment, the analysis of the above result is trivial, but one minor observation is that for accuracy less than 85, the music played by the novice user is hardly recognizable as the original melodies.

The feedback questionnaires include the following 6 questions and the answers have a scale from one to seven.

1. What do you think of the visual GUI of the MusicPattern?
2. Were you able to understand the meaning of timing axis (vertical axis)?
3. Were you able to understand the meaning of pitch axis (horizontal axis)?
4. Were you able to understand the basic musical concept from the tabletop canvas?
5. Was the finger operation intuitive (select, cancel, drag, multi-touch)?

6. Was the replay game fun?

7. [Open question] Are there any suggestions to the game design?

The numerical statistics of the participants' answers are plotted in the figure below:

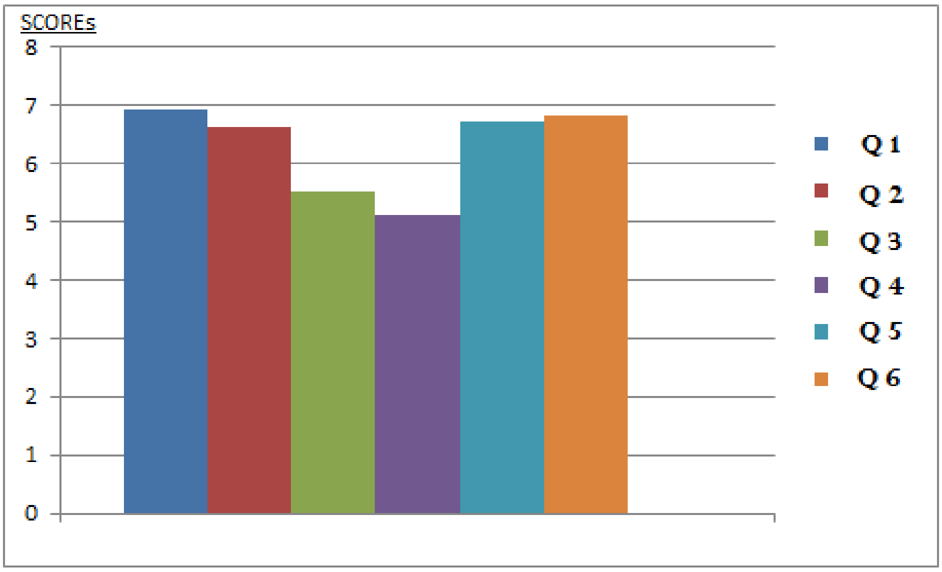


Figure5.5: questionnaire feedback: Average

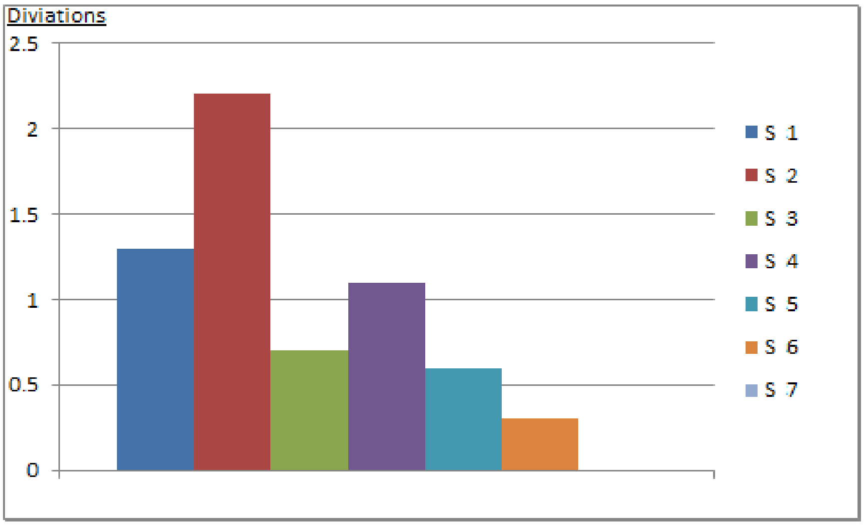


Figure5.6: questionnaire feedback: Standard deviation

From the questionnaire, the feedback was quite positive, especially for the GUI design and visual game design questions. For question #3, most layman users were able to understand the meaning of pitch axis as most of them are quite familiar with the pitches. In contrast, for question #4, they are not able to fully appreciate the abstract musical concepts within the duration of the short experiment. But it did give the layman user an overall feeling of these concepts. The game design receives the best feedback in question #6, probably due to its entertainment nature as well as the fun of participation. In question #2 asking about the meaning of timing axis, the standard deviation is quite high. This reflects that individuals have different perceptions and understandings in the time related musical features and the time taken by each individual to appreciate the use of time axis is quite different.

Other than the six questions the novice users also feedback that the MusicPatterns should also include text instructions on the application menus as well as the canvas. Although this is a very good suggestion, the reason why the application does not include text descriptions of the functions is because the project is trying to promote the intuitiveness of use. The testing version of the application aims to see how much a layman users can understand the functions purely by first impressions without text instructions. It is also useful to have deeper observations of the users' behaviors as how this GUI can be fast accepted and how efficient the finger operations are on the table top surface when they are presented with the new MusicPatterns UI.

Another point collected from layman's feedback is that under the replay game mode, it was very difficult to catch the curves, which always reduced the overall result. As the main objective of this research design is not about the game, fewer discussions are made in this area. But it is believed that this problem can be resolved, provided that the detection code includes a tolerance level where the finger touch area with the deviation of one or two pitches can be accepted as a correct audio pitch output.

5.4.3 For Experienced group

The Experiment and Discussions

The NAFA group was given a group task of composing three 37-seconds standard music melodies using the MusicPattern application. The music creation was under the tube mode and it was not an individual task but a collaborative one. The 5 experienced composers decided to create three music melodies: a Cannon D rock with 3 musical instruments; a pop music with 4 musical instruments and two imported musical instrument channels; and a solo melody which was the original work created by one of the participant.

The participants first divided the task into two parts. The first task was to prepare the tubes needed. Each tube contained an average of 7 pitches. For each musical instrument, the channel of the chorus instrument is normally less than ten, so the number of tubes is around ten; and for the percussion instrument, there are normally 7-8 types of sound, so we need 7-8 notations on tabletop for finger to perform drum effect. For the solo channel, they intended to use a smaller number of tubes with relatively more pitches inside each tube. This was to enable the fast operations and the pitch bending effects that is always required by the solos. The solo tube requires more advanced skills of operation in real time music creation due to its faster speed of operation. The real operation is difficult because the bending effect is pointing to different directions and the magnitude is not easy to control in real music playing. The same issue is observed for practicing most of musical instruments. But the rest of the channels worked perfectly well. Most of composing times were used to modify the solo patterns.

The three pieces of work were produced within 3 hours. A sample screenshot for the pattern created is shown in the figure below:

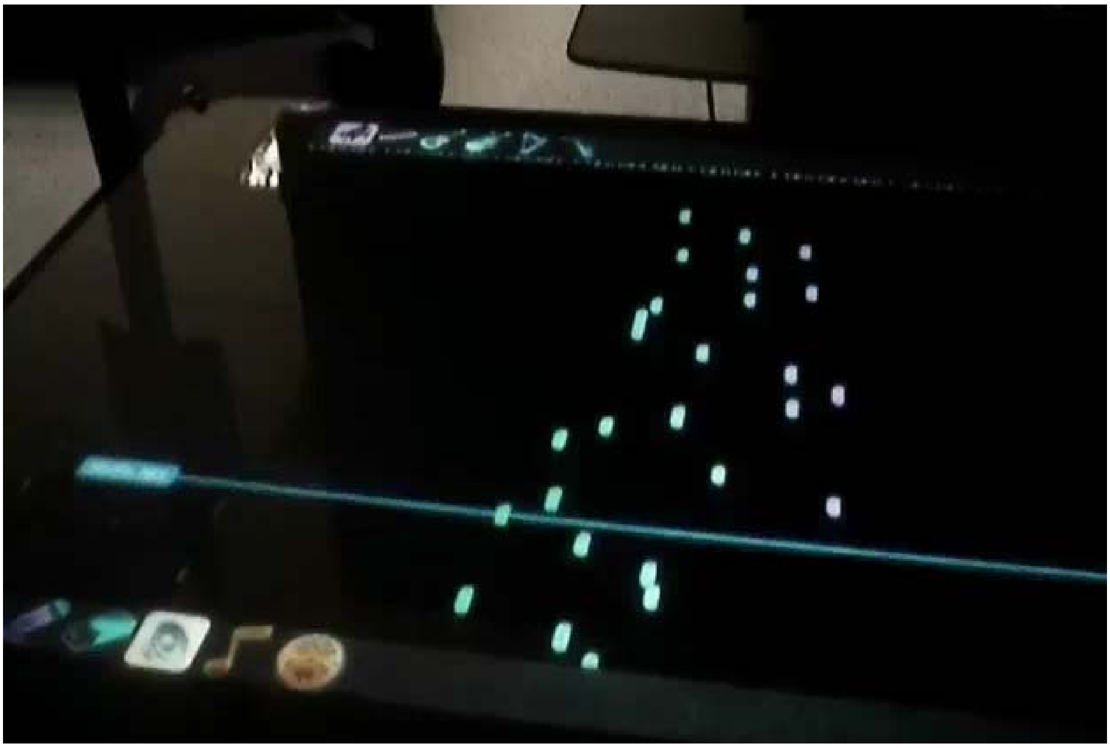


Figure5.7: Cannon segment video screenshot

The questionnaire for the experienced users has eight questions and the answers have a scale from one to seven except for the last open questions asking for more feedbacks. The questions are listed below:

1. What do you think of the idea to represent music using visual pattern?
2. Are the main musical features in the musical melody composition included in the pitch time table two-dimensional layout in this system design?
3. Has the visual layout actually helped in the observations of musical feature definitions?
4. Is the pitch bending function easy to use?
5. How do you think about the collaboration when doing a group composition task?

6. What is the overall efficiency of composing a real-time piece of music?

7. What is the overall efficiency of modifying the music off-line by modifying patterns?

8. Is there any further suggestions?

The results are given in the figure below:

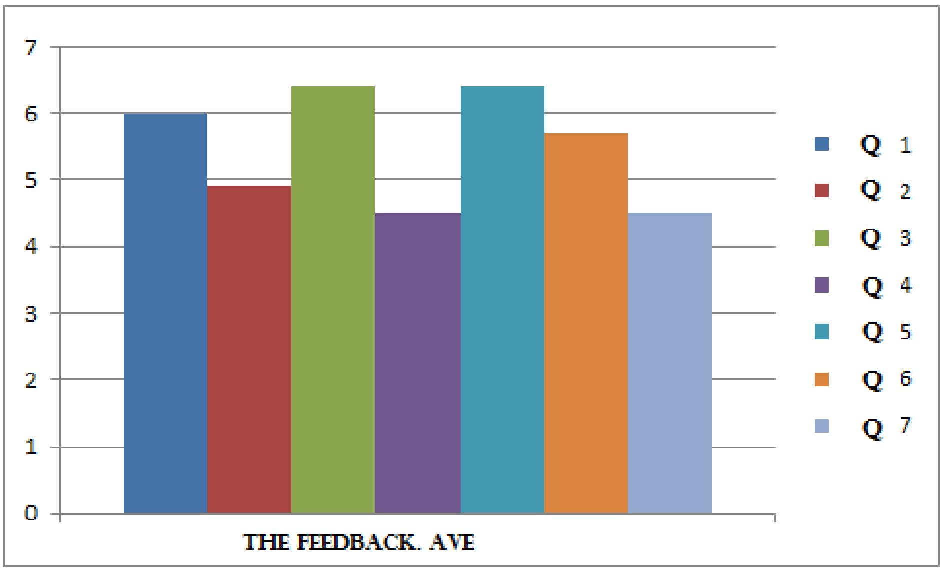


Figure5.8 The average of the experienced feedback questions

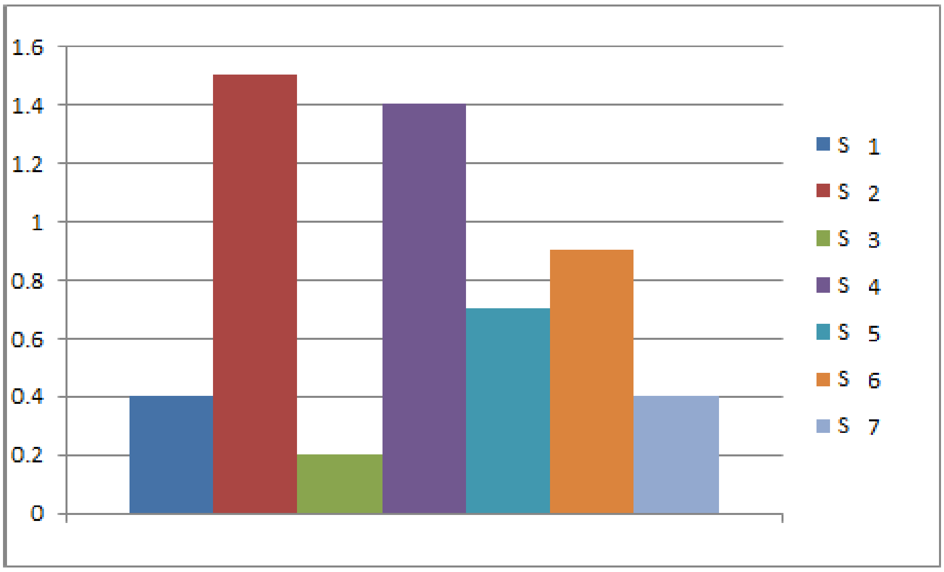


Figure5.9 The standard deviation of the experienced feedback questions

The feedback for question #2 “Are the main musical features in the musical melody composition included in the pitch time table two-dimensional layout?” reflects a critical issue affecting the music composition. The volume control feature is not included in the project design. And the major part of the future works will be dedicated to this area in order to create a platform that can fully represent most of the important features. The dimension is one of the major limitations when we are trying to include more features. To resolve this issue, the thickness of the line and the colors can be added to increase the overall data handled up to four dimensions. For question #4 “Is the pitch bending finger operations easy to use?”, the tube mode creatively introduced the pitch moving functionalities where the pitches are not confined inside the tubes, and that makes the pitch bending and tremolo effects possible for operation[41]. However, the finger control is difficult in practice due to the sensitivity of the hardware laser as well as the speed of real time solos. To resolve this issue, the first thing to do is to upgrade the laser system. However, as the solo speed increases, the composer still has to face a real time playing difficulty under the current tube design. It is proposed to create a special tube designed just for solos, where the pitch balls are queued inside the tube but only the left most balls can be played. Once the left most pitch is released, it will be popped out of the tube, and the rest of the pitch ball queue will be moved towards the left. By this recording method, the solo playing are much better controlled as the location of the pitch is fixed and the user only needs to tab, push, and release on one pitch ball.

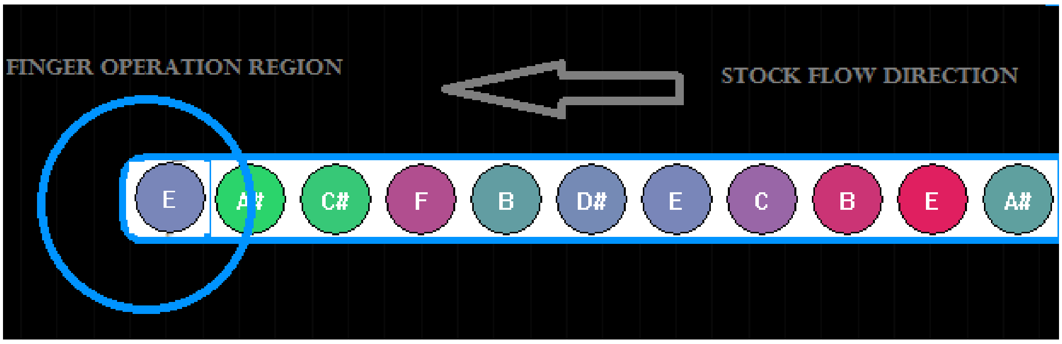


Figure5.10. Tube-Stock mode

The first-in first-out pitch balls in tube mode was integrated into the latest version of software.

The off-line editing is performed after the real-time music playing. The modifying functionalities are quite poor in this prototype design, which has been reflected in answers to the question #7. In the answers to question #8 for future work suggestions, the same problem was also mentioned.

The users' feedback indicates that the major merits that the students found from Music Pattern GUI are: 1) the canvas y-axis can easily help team mates to locate any time location where discussion or disagreement can be visually seen and analyzed; 2) the finger touch is useful in rhythm control under synchronized composing while in text notation it is something difficult [42]; and 3) the user found this collaborative work fun and interesting in the cooperation and interaction among multi-instruments.

However, the group also pointed out certain limitations of design. The most important one was the lack of the volume control, without which the information that carries a significant amount of emotional expressions is missing. They also suggested use the thickness of the line pattern to represent volume variations. Moreover, they also felt that under finger editing mode, there should be more editing tools to increase the editing efficiency. An example would be a tool box with functions to copy and paste paragraphs, which would be quite helpful when producing the repeated parts of melodies. Another limitation they pointed out is pertaining to the poor quality of the audio output, which is because the current version of design only makes use of Windows' default wave table for MIDI instruction. To use a better audio tone synthesizer will definitely improve the audio quality which is beyond the design concept of Music Patterns. Thirdly, they mentioned the “tolerance problem” that is due to the over-sensitivity of finger tracking system, which sometimes caused the floating finger or the shadow of fingers to produce unwanted output. Furthermore, they also suggested the curvy tube designs to make the shape of the tube more adapted to the natural shape of the hand and fingers, thus enhancing the overall composing experience.

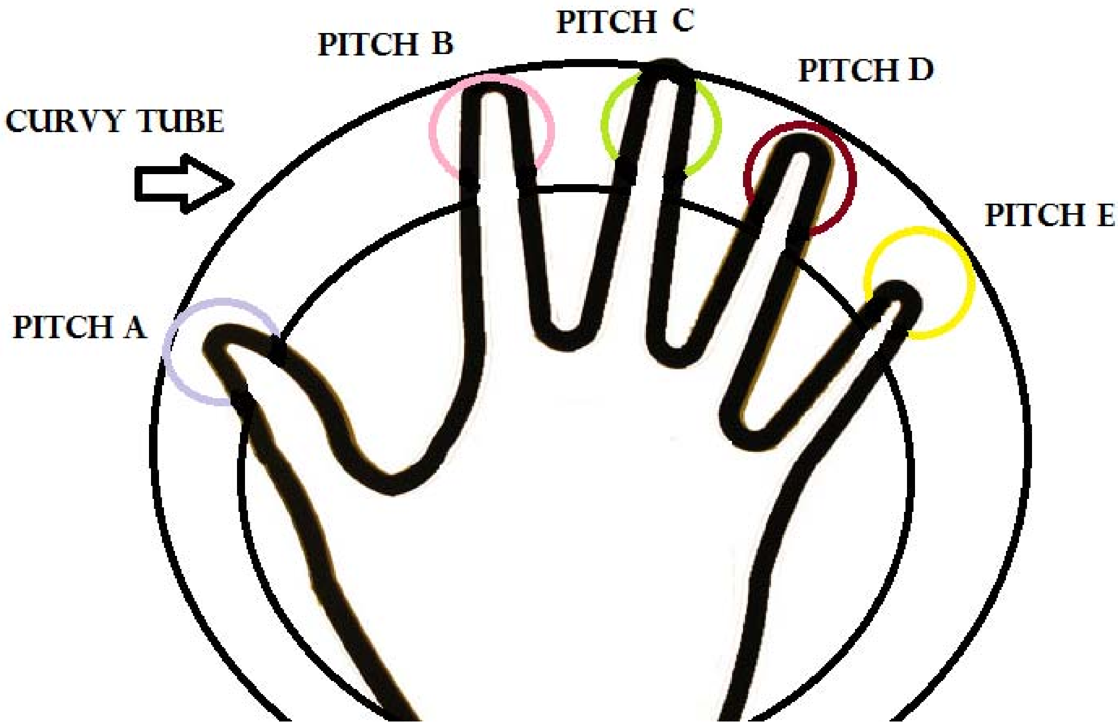


Figure 5.11 Curvy tube demo.

Chapter Six Conclusion and Future Works

This project proposed and created a tangible musical interface which aims to promote the easy and fun operational concept in music melody composing tasks. The design can also be exploited for potential educational purpose of music training for the layman users, as the interface is scientifically designed to represent most of the musical terminologies in visual forms. Both the software interface and the finger operations support by the hardware design aim to lower the common music composing barriers in order to enrich the music playing experience for everyone. On one hand, this design makes use of a pitch-time based 2D graphic for music representation to provide users with intuitive and visual impact on melody structure. On the other hand, the multi-touch tabletop system further supports direct finger operations to facilitate the music editing and creation in practice. To let the users see and touch while hearing music melodies is what we believe an intuitive manner for them to perceive and explore how the music is created. In this paper, we presented the literature review, the game design, the implementation, and the early users' experiences. The design limitation and future works will also be discussed in the conclusion session.

The future work is planned based on the on-going users' studies. The current feedback requires the system to further support volume control and editing tool box. Even though the current MusicPattern presented here may not be a professional music composing platform and does not support advanced controls for complex composition, we argue it is one of the most intuitive and meaningful designs among the emerging tabletop applications. The project only demonstrated the musical tabletop software design framework and included the fundamental musical features and concepts. In order to create an interactive musical platform or even exploit its commercial values, the necessary future works will be focusing on the following areas:

1. The Volume Representation

As discussed in the user study chapter, the volume of the pitch lines is critical for emotional communications. To enable the volume control, a suggested solution is to use the thickness of lines or size of the pitch balls to represent the volume. However, if doing so, the space between neighboring pitches in the pitch axis layout will be readjusted, which requires additional coding efforts.

2. More Useful Editing Tool Box Functions

One of the advantages possessed by the professional text-based composing software is that they have very powerful editing tools. However, in the tangible tabletop, it is not easy to design all those functionalities within a short project timeline, as the implementation of each of the functional buttons requires deep understanding of musical concepts, and this in turn requires the designer to have a strong musical composing background. In this engineering design project, the author did not focus on this part.

3. Audio Voice Input

While using visual representations and finger controls has made the musical composing procedures intuitive and accessible to even less musically trained users, what can be even simpler is to use human voice input in musical melody generations and games [12]. The pitch recognition work has already been completed for years [45], but the difficult issue is the integration with the MusicPattern platform as well as the real-time matching issue. However, the author is confident that this is possible to achieve.

4. Tone Editor

   Another way of utilizing the tangible interface is the music tone constructions. The musical instruments are characterized by the formant in frequency spectrum[45]. To visually present that curve for users to drag and create multi-tone formants will be an interesting research area. From the tone editor, the users would be able to learn how musical instrument are different from each other and they are also possible to try and create some new tones that has never occurred in the history of musicology.

BIBLIOGRAPHY

[1] Wyeth, Peta, “Understanding engagement with tangible user interfaces” : Designing for Habitus and Habitat, OZCHI'08, p 331-334, 2008, Proceedings of the 20th Australasian Conference

[2] Roger B. Dannenberg, “Interactive Visual Music: A Personal Perspective”. Computer Music Journal, Volume 29, Number 4, Winter 2005, pp. 25-35

[3] Seth Price, “Early Video Game Soundtracks” Adapted from an article published in In Magazine, 2002

[4] Jonathan B. Prince, “The integration of pitch and time in music perception” Doctor thesis of Philosophy, 2009 Department of Psychology, University of Toronto.

[5] Cheok, A., Yang, X., Ying, Z., Billinghurst, M., and Kato, H. 2002. Touch-Space: Mixed Reality Game Space Based on Ubiquitous, Tangible, and Social Computing. Personal Ubiquitous Comput. 6, 5-6 (Jan. 2002), 330-442.

[6] Lynch, Kevin (2009-11-06). "Game review: Band Hero". The Mirror. Retrieved 2009-12-06.

[7] Morris, M.R., Piper, A.M., Cassanego, A., Huang, A., Paepcke, A., and Winograd, T. Mediating Group Dynamics through Tabletop Interface Design. IEEE Computer Graphics and Applications, Sept/Oct 2006, 65-73

[8] Kaltenbrunner, Günter Geiger, Nadine Alber, The reacTable: A Collaborative Musical Instrument, in Proceedings of the 15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises. 2006, IEEE Computer Society.

[9] Pedersen, E.W., K. Hornb, and #230, mixiTUI: a tangible sequencer for electronic live performances, in Proceedings of the 3rd International Conference on Tangible and Embedded Interaction. 2009, ACM: Cambridge, United Kingdom.

[10] Jord, Sergi Jordà, The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces, in Proceedings of the 1st international conference on Tangible and embedded interaction. 2007, ACM: Baton Rouge, Louisiana.

[11] Kaltenbrunner, M. and R. Bencina, reacTIVision: a computer-vision framework for table-based tangible interaction, in Proceedings of the 1st international conference on Tangible and embedded interaction. 2007, ACM: Baton Rouge, Louisiana.

[12] Levin, G. and Z. Lieberman, In-situ speech visualization in real-time interactive installation and performance, in Proceedings of the 3rd international symposium on Non-photorealistic animation and rendering. 2004, ACM: Annecy, France.

[13] Hiraga, R. and N. Matsuda, Visualization of music performance as an aid to listener's comprehension, in Proceedings of the working conference on Advanced visual interfaces. 2004, ACM: Gallipoli, Italy.

[14] Sandlers, S. http://sethsandler.com/audiotouch/. Referenced August 28, 2009.

[15] Kaltenbrunner, M. & Bovermann, T. & Bencina, R. & Costanza, E. "TUIO - A Protocol for Table Based Tangible User Interfaces", Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation (GW 2005), Vannes (France)

[16] Wright, M., Freed, A., Momeni A.: "OpenSound Control: State of the Art 2003". Proceedings of the 3rd Conference on New Instruments for Musical Expression (NIME 03), Montreal, Canada, 2003.

[17] [Q2] Schouten, J. F. (1968). "The Perception of Timbre". In Reports of the 6th International Congress on Acoustics, Tokyo, GP-6-2, 6 vols., edited by Y. Kohasi, p. 42. Tokyo: Maruzen; Amsterdam: Elsevier.

[18] Schouten, J. F. (1968). "The Perception of Timbre". In Reports of the 6th International Congress on Acoustics, Tokyo, GP-6-2, 6 vols., edited by Y. Kohasi, p. 42. Tokyo: Maruzen; Amsterdam: Elsevier.

[19] Review and Pictures of Apple iPod Nano MP3 Player (MA099LL/A,MA004LL/A, MA107LL/A,MA005LL/A),http://www.reviewshop.com/iPod_Nano/iPod_Nano_Review.html

[20] Chia Shen Mitsubishi Electr. Res. Labs., Cambridge, MA, USA. “Multi-user interface and interactions on direct-touch horizontal surfaces: collaborative tabletop research at MERL” Horizontal Interactive Human-Computer Systems, 2006. TableTop 2006.

[21] Introduction to OSC; The OSC Specifications; http://opensoundcontrol.org/introduction- osc. Upadated on 2010.Sep 17.

[22] John Fuller, Univesitat Pompeu Fabra in Barcelona, Spain “How the reactable Works”- Electronics Deconstructed Videos, p2-p11.

[23] NUI Group Community Forums, “ C/C++ (CCV/OpenCV/OpenGL/OpenFrameworks)” “Getting Started with CCV”, Update Feb 2010.

[24] Loren Heiny, “First successful experiments with a multi-touch box” Incremental Blogger » Blog Archive » Multi-touch videos catch on ; May 11, 2008

[25] Image credit for Figure 2.3, The reacTable interface: Martin Kaltenbrunner, Günter Geiger, Marcos Alonso, http://www.ycam.jp/en/art/2008/10/minimum-interface.html

[26] Esben Warming Pedersen University of Copenhagen, Copenhagen S, Denmark Kasper Hornbæk University of Copenhagen, Copenhagen S, Denmark, “mixiTUI: a tangible sequencer for electronic live performances” TEI '09 Proceedings of the 3rd International Conference on Tangible and Embedded Interaction.

[27] Image citation: Figure 2.4, User Manipulating spectrograms on Scrapple, http://www.flong.com/projects/scrapple/

[28] Berry, R. (ATR Media Inf. Sci. Labs., Kyoto, Japan); Makino, M.; Hikawa, N.; Suzuki, M. “The augmented composer project: the music table”,Proceedings Second IEEE and ACM International Symposium on Mixed and Augmented Reality, p 338-9, 2003

[29] The Bush Telegraph: networked cooperative music-making
Berry, R. (ATR Media Inf. Sci. Labs., Kyoto, Japan); Makino, M.; Hikawa, N.; Suzuki, M. Source: Entertainment Computing - ICEC 2004. Third International Conference. Proceedings (Lecture Notes in Comput. Sci. Vol.3166), p 120-3, 2004

[30] Image citation: Figure 2.5, Interface of AudioTouch, http://sethsandler.com/software/video-featuring-more-multitouch-software/

[31] The use of a table driver for individualized design of computer-based instruction materials in music theory,Arenson, M. (Univ. of Delaware, Newark, DE, USA) Source: 1982 Conference Proceedings of the Association for the Development of Computer-Based Instructional Systems, p 228-30, 1982

[32] Tracking music from media using reduced pitch class profile features
Unal, E. (Ulusal Elektron. ve Kriptoloji Arajtirma Enstittisti, TUBITAK, Ankara, Turkey); Akin, A.A.; Kanak, A.; Dogan, M.U. Source: 2009 IEEE 17th Signal Processing and Communications Applications Conference (SIU), p 373-6, 2009 Language: Turkish

[33] Image citation: Figure 2.6, Users collaborating locally on mixiTouch, http://journal.mycom.co.jp/column/svalley/257/index.html

[34] Tan, S.Y.P. (Sch. of Comput., Nat. Univ. of Singapore, Singapore, Singapore); Zhijia Hu; Koh, A.Y.L.; Felicia; Shengdong Zhao Source: Proceedings of the SPIE - The International "MusicFlow: an interactive music composition system",Society for Optical Engineering, v 7744, p 774408 (7 pp.), 2010

[35] CHRIS WILLMAN; "Music Features Archive" October 18, 2010 , Newyork Magzine p132-133

[36] CAMP: computer-aided music processing,Nieberle, R.C. (Tech. Univ., Berlin, Germany); Koschorrek, S.; Kosentzy, L.; Freericks, M. Source: Computer Music Journal, v 15, n 2, p 33-40, Summer 1991

[37] Kawanobe, M. (Sch. of Manage. & Inf. Sci., SANNO Univ., Kanagawa, Japan), "Affective mapping between music and color combination" Source: 2005 5th International Conference on Information Communications and Signal Processing (IEEE Cat. No. 05EX1118C), p 5 pp., 2005

[38] Hayashi, S. (NTT, Tokyo, Japan); Honda, M.; Kitawaki, N. "Adaptive predictive coding scheme and its real time processing at 64Kb/s for wide-band speech and music signals" Source: Electrical Communication Laboratories Technical Journal, v 32, n 11, pt.1, p 2327-38, 1983

[39] Wang Yue (Speech & lyinguage Process. Group, Northwestern Polytech. Univ., Xi'an, China); Xie Lei; Yang Yu-lian "Adaptive whitening for real-time music beat tracking" Source: Application Research of Computers, v 26, n 5, p 1676-8, 1684, May 2009

[40] Mozer, M.C. (Dept. of Comput. Sci., Colorado Univ., Boulder, CO, USA) “Neural network music composition by prediction: exploring the benefits of psychoacoustic constraints and multi-scale processing” ,Connection Science, v 6, n 2-3, p 247-80, 1994

[41] Stavness, I. (HCT Lab., British Columbia Univ., Vancouver, BC, Canada); Gluck, J.; Vilhan, L.; Fels, S. “The MUSICtable: a map-based ubiquitous system for social interaction with a digital music collection” Entertainment Computing - ICEC 2005. 4th International Conference. Proceedings (Lecture Notes in Computer Science Vol. 3711), p 291-302, 2005

[42] Shmulevich, I. (Signal Process. Lab., Tampere Univ. of Technol., Finland); Povel, D.-J. “Rhythm complexity measures for music pattern recognition”,1998 IEEE Second Workshop on Multimedia Signal Processing (Cat. No.98EX175), p 167-72, 1998

[43] Ukkonen, E. (Dept. of Comput. Sci., Helsinki Univ., Finland); Lemstrom, K.; Makinen, V. “Sweepline the music! [pattern matching]”, Computer science in perspective. Essays dedicated to Thomas Ottmann, p 330-42, 2003

[44] Shmulevich, I. (Signal Process. Lab., Tampere Univ. of Technol., Finland); Yli-Harja, O.; Coyle, E.; Povel, D.-J.; Lemstrom, K. “Perceptual issues in music pattern recognition: complexity of rhythm and key finding”, Computers and the Humanities, v 35, n 1, p 23-35, Feb. 2001

[45] Mont-Reynaud, B. (Dept. of Music, Stanford Univ., CA, USA); Gresset, E. PRISM: “Pattern Recognition In Sound and Music”, ICMC Glasgow 1990. Proceedings, p 153-5, 1990

[46] Chord (music), Wikipedia, the free encyclopedia, Section: Triads; http://en.wikipedia.org/wiki/Chord_%28music%29