Digiti Sonus: Advanced Interactive Fingerprint Sonification Using Visual Feature Analysis

Yoon Chung Han
Media Arts and Technology
University of California
Santa Barbara, USA
yoon@mat.ucsb.edu

Byeong-jun Han
School of Electrical Engineering
Korea University
Seoul, Korea
hbj1147@korea.ac.kr

Matthew Wright
Media Arts and Technology
University of California
Santa Barbara, USA
matt@create.ucsb.edu

ABSTRACT

This paper presents a framework that transforms fingerprint patterns into audio. We describe *Digiti Sonus*, an interactive installation performing fingerprint sonification and visualization, including novel techniques for representing user-intended fingerprint expression as audio parameters. In order to enable personalized sonification and broaden timbre of sound, the installation employs sound synthesis based on various visual feature analyses such as minutiae extraction, area, angle, and push pressure of fingerprints. The sonification results are discussed and the diverse timbres of sound retrieved from different fingerprints are compared.

Keywords

Fingerprint, Fingerprint sonification, interactive sonification, sound synthesis, biometric data

1. INTRODUCTION

Fingerprints are one of the most unique biometric patterns on the human body. Due to the distinct patterns and various visual characteristics of the patterns, they have been useful for identification and security, typically based on extraction of visual features. We believe that sonification can serve as an effective technique for the representation of complex information, due to the auditory system's ability to perceive stimuli at a wide spatial cover and its inclination to perceive spatial patterns in sonic input. In addition this work aims to provide a way for visually impaired people (as well as those with normal vision) to experience fingerprints, potentially even learning to recognize specific people's fingerprints by sound. We believe this is the first published work on fingerprint sonification in either the scientific or artistic domain. Our initial development is therefore, artistic and experimental because it can closely bridge the relation between body, sound, visuals, and interactivity in novel ways.

2. BACKGROUND

We introduced *Digiti Sonus* at the 2012 ACM Multimedia conference [1], describing sonification based on visual feature extraction from a scanned fingerprint image. The fingerprint's ridges and valleys are acquired using skeletonization to eliminate pressure-specific information. Skeletonization also normalizes the ridge and valley portions to acquire the ridge skeletons. From this we extract the *minutiae*, which are the major features of a fingerprint, such as ridge endings or ridge bifurcations. Minutiae are scanned in one direction, the scanned fingerprint image is transformed into a magnitude spectrogram, and audio filters are applied to the spectrogram. In this step, audio filters can be anything such as traditional low-/high-/band-pass filters or user-defined function. In order to specify the characteristics of audio filters, we transformed the properties of the detected minutiae into such as position, direction, and type.

Key insights over the past year have led to improvements in every aspect of the system. The most significant improvement is in the diversity of output sound. In previous research, the sound results had a limited range of timbre and frequency, thus users had difficulty understanding the uniqueness of each fingerprint. Furthermore, one-directional scanning was too limited and resulted in an overly uniform range of sounds due to using only a single animation direction. We decided to explore more dynamic ways of reading through fingerprint patterns. Experience installing this work in a gallery and receiving audience feedback led to change the installation setup as well. The previous prototype contained only one projector and the fingerprint scanner, however, due to the dynamic 3D visuals and audio in new version, we changed to three projectors to show three different perspectives of a fingerprint as well as adding an interactive touch screen alongside the fingerprint scanner.

Among these problems, we address the diversity of sound as the most important issue since audience should be clearly aware of distinction between different fingerprints. This paper focuses on enlarging the range of timbre in sound synthesis. Section 3 describes related works, and Section 4 discusses conceptual approaches and visual feature extraction of fingerprints. In Section 5, the sonification and visualization of fingerprints are deeply described and the implementation of those approaches is the subject of Section 6. In next section, we evaluate the final results, and in the final section we state our conclusions and the future extension of Digiti Sonus.

3. RELATED WORK

Although we could not find any previous research on fingerprint sonification, there is a deep body of work on biometric fingerprint pattern recognition [2]. It is both an active academic research topic and also a practical problem with commercial/industrial applications and implementations. On the other hand, there are many examples of sonification of other body data or body patterns, which influenced our research. Worgan [3] sonified faces using multiple sound synthesis techniques such as additive, wave terrain, and frequency modulation to represent facial data, which is the visual focal point for human-to-human interaction. Although the mapping of each facial feature (eye, mouth, eye color, face color, etc.) in this research is novel, the mapping is not strongly effective to deliver the visual data. It is very hard to make conclusions about the effectiveness of any mapping without conducting extensive user studies. The mapping should describe not only

NIME'13, May 27-30, 2013, KAIST, Daejeon, Korea.

Figure 1. Extracted minutiae of fingerprints through image skeletonization and minutiae analysis.

Figure 2. Various captured fingerprints.

differences of faces but also deliver appropriate sound psychologically. This issue is usually the most challenging part in sonification, and this research has led us to deeply consider the mapping method of sonification. Daito Manabe [4] created "Face Visualizer," which controls music with his face. Electrical sensors attached to the surface of his face detected the electrical pulses that control muscle movements. His work transformed body into sound directly, and could make live music performance happen in real time, which is the most beneficial part. The real-time performance aspect of this work influenced our project.

Although most sonification of body data may refer to sonified body movements or dynamic gestures over time (from which perspective almost every NIME could be considered to perform some form of body data sonification), there are also examples of sonifying inner/invisible body data, such as brain activity (e.g., brainwaves, biorhythm, emotions, or sleep). Weinberg and Thatcher [5] developed the sonification of neural activity in their "Fish and Chips" and "BrainWaves" projects. Angel [6] also applied an artistic approach to the transformation of brainwave data. In her installation and performance, EMG (Electromyography) and EEG (electroencephalography) were used to explore the interaction between the human body and real-time media. Brainwaves result in dynamic datasets and invite interactive sonification.

This prior work has inspired our research to explore more dynamic and interactive approaches to fingerprint sonification. Although fingerprints are fundamentally static data, users can control and modify the presentation of their fingerprints; therefore our current implementation attempts to detect these interactive parameters and incorporate them in the sonic results.

4. DIGITI SONUS

4.1 Fingerprints as Sonic Identities

Digiti Sonus is an interactive audiovisual art installation based on fingerprint sonification. Transforming fingerprints' unique patterns into sonic results allows the audience to experience the discovery of sensory identities; from an artistic perspective, finding their own sound identities through the fingerprints gives a unique experience to the audience.

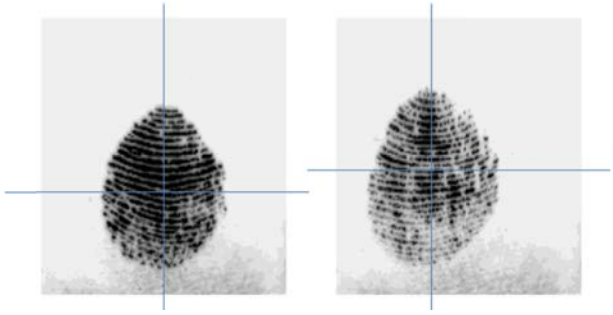


Figure 3. Two images captured from a fingerprint with different push pressure.

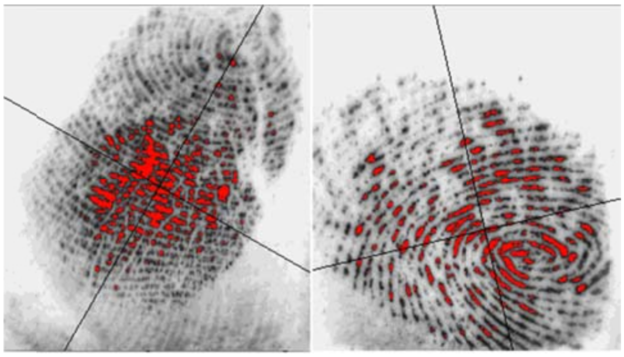


Figure 4. Different angles of two fingerprints.

We treat the distinct visual features of fingerprints as an open musical score whose performance can be executed in diverse ways. By controlling the starting point of animated visuals, the predetermined musical notes can be reorganized in different orders and duration.

Fingerprints provide rich and detailed input data suitable for expression via musical parameters. We believe that delivering these data in immersive, sensory way helps make understanding the complicated patterns of fingerprint easier and simpler. Visitors to the installation can "perform" musical sound (albeit in an extremely limited way; see next section) without any difficulty, providing input that results in musical sound.

This artwork is designed as either a gallery installation or a software/application. In a gallery setting, the audience can touch a fingerprint sensor and experience the audiovisual output in real time in the immersive environment. In a software application, users can input their fingerprints via their own sensor and observe the sonification process, supporting identification in aural attention or educational expression for children or students.

4.2 Performance Practice

Digiti Sonus was exhibited in Geumcheon Seoul Art Space as a part of Media City Seoul 2012 in September 2012. During the exhibition, we observed some unanticipated audiovisual results and unique performances from the audience. Even though every fingerprint is essentially a static dataset, audience members were able to vary their audiovisual results by presenting their fingers to the scanner in different ways: the pressure, orientation, and size of the fingerprint input all greatly affect the resulting image. Thus, one person could perform various sounds with only one finger. This observation led us to examine how to expand the diversity of audio results, and this research can be expanded to the sonification of dynamic spatial data. We analyzed the patterns and dynamic features of a static fingerprint data based on previous visual feature extraction results.

4.3 Visual Feature Extraction

We extracted five distinct visual features that can be analyzed from single fingerprint such as the one shown in Figure 1.

4.3.1 Position of Minutiae

The minutiae, the ridge characteristics of fingerprints, can be extracted throughout the process of image skeletonization as described in [1] and shown in Figure 1. The two most prominent local ridge characteristics, called minutiae, are ridge ending and ridge bifurcation [7]. A ridge ending is defined as the point where a ridge ends abruptly. A ridge bifurcation is defined as the point where a ridge forks or diverges into branch ridges. A typical fingerprint contains about 40–100 minutiae. Usually, automatic fingerprint matching depends on the comparison of these minutiae and their relationships to make a personal identification [8].

Thus, the positions of minutiae are the key information to recognize the distinction of fingerprints, and widely used for identification. We therefore chose this unique characteristic to be the most significant factor for sonification. For the sake of simplicity, we ignore the type of each minutia.

4.3.2 Area, Pixel Range, and Push Pressure of Fingerprint

Figure 2 shows several results of placing fingers on the fingerprint sensor screen with varying position, orientation, and pressure. Some are larger and fill the screen fully, while others only cover a small portion of the screen. In addition, some depict bright images, whereas others show dark images. Figure 3 shows two images from the same finger, but with the left one stronger and the right one larger.

Our analysis first selects only those pixels over a given brightness threshold; for example in Figure 4 the red dots indicate the selected (sufficiently bright) pixels. The “position” of the fingerprint as a whole is the 2D centroid of the locations of the pixels over the threshold. (In figures 3, 4, and 9 the center of the cross superimposed over each fingerprint represents the centroid.) The “area” of the fingerprint is simply the number of pixels above the threshold. Finally, the push pressure is the mean brightness of the pixels over the threshold.

4.3.3 Angle of Fingerprint

Users can apply their fingers to the fingerprint sensor at any angle. Usually, users touch the sensor near 0° (perfectly vertical orientation), however users often slightly rotate the finger, generally depending on the particular finger (thumb, index, etc.) and hand (left versus right). Our analysis of fingerprint angle is based on PCA (Principal Component Analysis) [9]. Again, the input data are the x and y positions of the pixels above the brightness threshold. These data points are normalized by the mean of each dimension. Next, by computing eigenvalues and eigenvectors, the principal components are acquired. We regard the first eigenvector (principal component) as the direction of input fingerprint image. Since principal components are always orthogonal and our data is two-dimensional, the second component gives no additional information. Figure 3, 4 and 9 show the principal components of various fingerprints. Note that for a given physical angle of the user’s finger on the sensor, differences in push pressure, by determining which pixels are above the threshold, may also somewhat affect the detected angle.

5. INTERACTIVE FINGERPRINT SONIFICATION AND VISUALIZATION

In this section, we describe how we sonify and re-form the 2D fingerprint image into a 3D animated visualization based on the five distinct features described above. Also, we describe the mapping methods used to sonify the fingerprint dataset.

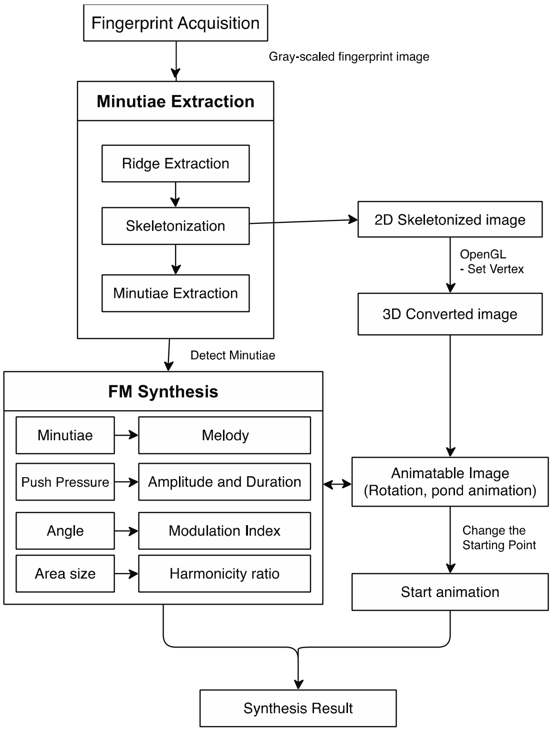


Figure 5. Overall fingerprint sonification and visualization process

In addition to the obvious form of interaction of having participants scan their fingerprints, Digiti Sonus also uses the metaphor of expanding circular ripples like those that would arise from dropping a stone into a pond. Alongside the fingerprint scanner is a visually inviting touch screen (described in Section 6.1); user touches on the screen interactively generate an outward-expanding ripple in the fingerprint image, centered on the touch point. We support only one ripple at a time; if a user retouches the screen before the previous ripple finishes then the old ripple disappears and a new one starts expanding from the new touch position. This ripple affects both the visualization (as shown in Figure 6) and the sonification (as described in Section 5.1.2).

5.1 Musical Expressions in Digiti Sonus

In early prototypes [1], we mapped the positions and number of minutiae into the frequency range of sound by regarding the whole fingerprint image as a magnitude spectrogram, with the x and y axes of the fingerprint image interpreted as frequency and time axes. In practice, this tended to create a very limited range of output timbres even though the minutiae of all fingerprints are distinctive. In order to broaden the diversity of timbre, we employed FM synthesis, with the five distinct fingerprint features mapped synthesis control parameters.

5.1.1 FM Synthesis

FM (Frequency Modulation) synthesis, discovered by Chowning in 1973 [10], is a way to alter the timbre of a simple waveform by modulating its frequency with another waveform. In the basic FM technique, a modulator oscillator modulates the frequency of a carrier oscillator [11]. One of the most significant benefits of FM synthesis is that a small number of input parameters easily control a large range of output sounds. This was a main reason we adopted FM, along with simplicity of implementation. Our Max/MSP implementation has five input parameters: fundamental/carrier frequency, amplitude,

Table 1. Transforming raw data and perceptual feature into FM Synthesis

Raw Data	Perceptual Feature	Variables of FM synthesis
Position of minutiae pixel	Position of minutiae	Melody
Average brightness of RoI pixels	Push pressure of finger	Overall Amplitude and duration
Angle of pixels above threshold selected pixels	Angle of fingerprint	Modulation index
Number of selected pixels above threshold	Area of fingerprint	Harmonicity ratio

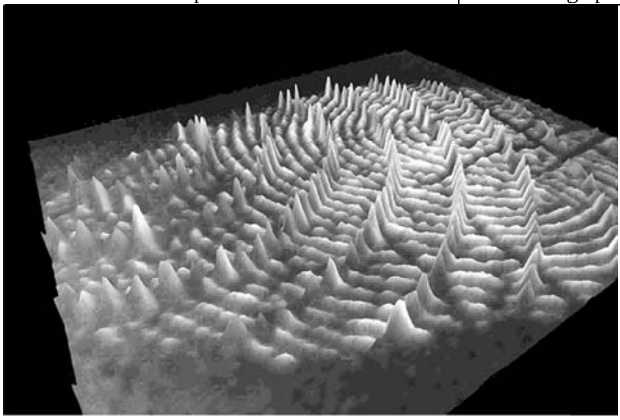


Figure 6. 3D Fingerprint Image

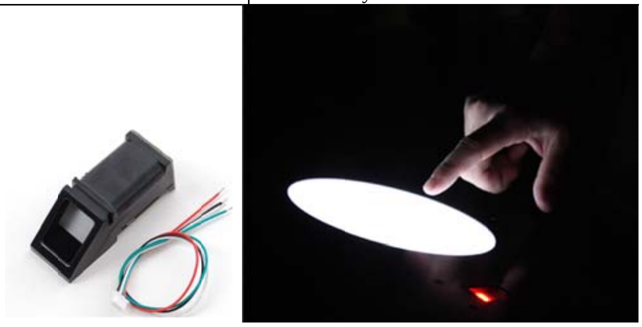


Figure 7. User input devices (a) fingerprint sensor ZFM-20, and (b) circular touch screen panel.

Figure 8. Installation Setup

modulation index (modulator amplitude over modulator frequency), and harmonicity ratio (modulator frequency over carrier frequency). An envelope controls the modulation index over the time of each note. The angle of the fingerprint affects the starting and ending points of the envelope that controls the modulation index. Thus, if the angle of fingerprint is higher, the angle of the envelope is also higher, which changes the modulator amplitude dramatically.

5.1.2 Mapping

The five visual features extracted from each fingerprint are mapped into control of FM synthesis as shown in Table 1.

Each minutia corresponds to a note. The interactive metaphorical stone-in-the-pond ripple triggers that note when the expanding circle reaches the minutia's 2D position. The note's frequency is determined by converting the minutia's 2D position into a single value using the technique of flattening the array of pixel positions: given a 2D (x, y) position where x goes from 0 to 255 (because of the fingerprint sensor's resolution), the corresponding 1D position is $256y+x$.

Fingerprint area maps to harmonicity ratio and angle maps to starting and ending point of one single envelope that controls modulation index. Push pressure maps intuitively to overall amplitude and duration, so that stronger imprints onto the fingerprint scanner result in stronger sounds.

Our mapping to FM synthesis is intuitive and has the property that obvious changes in the fingerprint come out as obvious changes to the sound. A larger and brighter fingerprint has higher magnitude and longer duration of sound. Higher number of minutiae results in a diverse pitch range and more total notes, often giving the percept of various melody lines. Users can be clearly aware of the relation between the area and brightness of their fingerprints and the harmonicity, magnitude, and melody range of the sound.

5.2 3D Visualization in Digiti Sonus

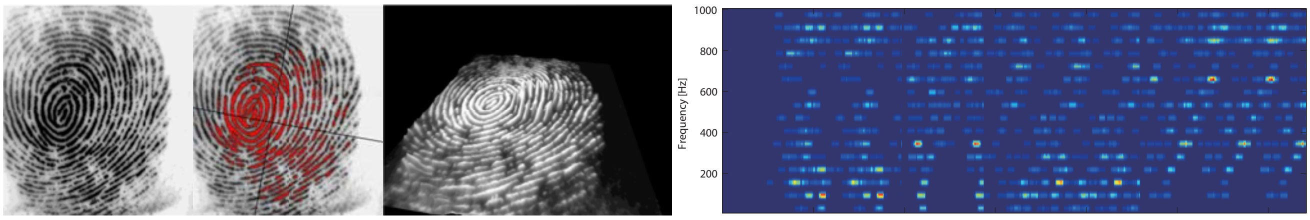
To increase user immersion, Digiti Sonus uses an animated 3D visualization. The captured fingerprint image appears on three projection screens. Since the user's fingerprint image is displayed in different perspectives, and ridges and valleys of the fingerprint image are divided by height of vertices, the user can observe his/her own fingerprint in from three diverse viewpoints.

Alongside the fingerprint sensor is a touch screen panel interface (described below) that allows users to indicate the starting point of audio playback. The indicated point is regarded as the center of a wave motion animated on the display screen. Whilst conducting this task, each wave on the screen is equivalent to audio playback, therefore, the user can feel the overall progress of fingerprint sonification more realistically. Moreover, the user may change the starting point of the wave motion (and hence the timing of the notes) by indicating another point on the touch screen interface. This property can give user interaction of both sound and display. Figure 6 depicts an instantaneous example of starting point selection and the result of sound playback and wave motion generation.

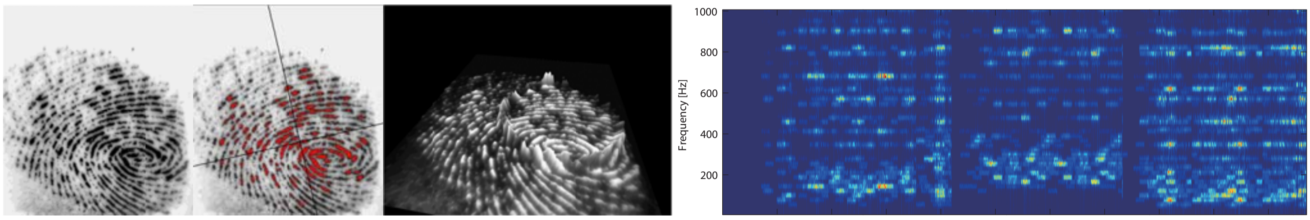
6. IMPLEMENTATION

6.1 Hardware System

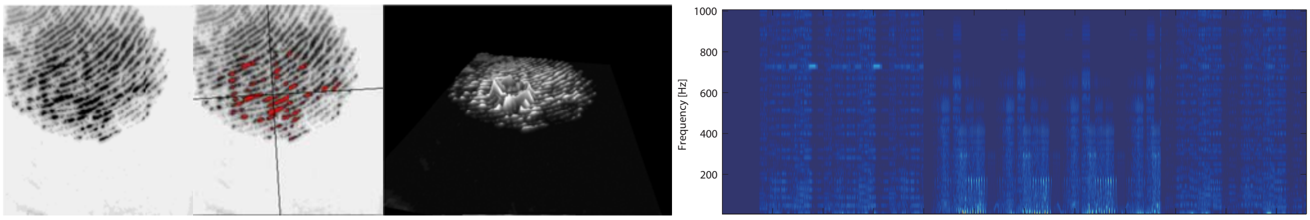
The current version of Digiti Sonus uses mostly the same hardware as the prototype [1], including PC system, and displays. The major differences between the current and prototype versions are as follows: 1) For the fingerprint sensor, we employed ZFM-20 [13] instead of Nitgen RS232 because of the ease of protocol implementation and circuit configuration and especially because of the increased speed with which it scans; 2) We installed multiple displays in order to show 3D fingerprint images in diverse perspectives, and; 3) A touch screen panel interface allows the user to determine the starting point of sound playback, as described in Section 5.2. The touch



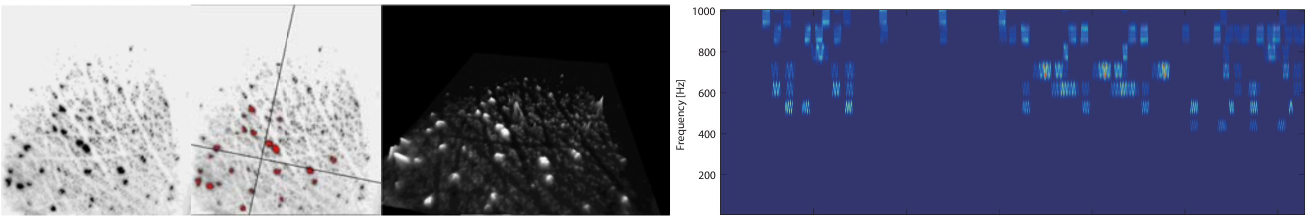
(a) Frequency range: 20~1026Hz, push pressure (amplitude): 0.8, area of fingerprint (harmonicity ratio): 1.0, duration: 5000.



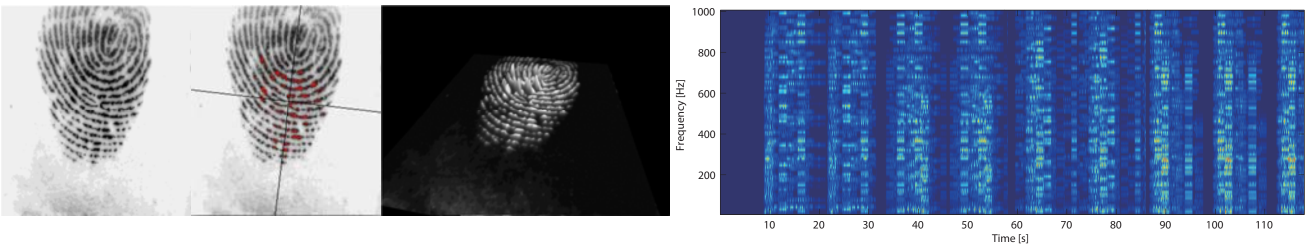
(b) Frequency range: 228~2000Hz, push pressure (amplitude): 0.57, area of fingerprint (harmonicity ratio): 0.5, duration: 3000.



(c) Frequency range: 20~950Hz, push pressure (amplitude): 0.3, area of fingerprint (harmonicity ratio): 0.2, duration: 1000.



(d) Frequency range: 430~1850Hz, push pressure (amplitude): 0.1, area of fingerprint (harmonicity ratio): 0.05, duration: 1000.



(e) Frequency range: 130~750Hz, push pressure (amplitude): 0.3, area of fingerprint (harmonicity ratio): 0.25, duration: 500.

Figure 9. Five input fingerprint Images, analysis, 3D displays, generated magnitude spectrograms from the input fingerprint images, and ranges of each parameters.

screen is of the type typically seen in touch monitors or touch-based interfaces, 10.4 inches wide and with a USB controller. Users can only touch a certain point, instead of swiping or dragging their fingers.

Considering the use of multiple displays and running multiple 3D fingerprint image displays in multi displays, we installed high performance VGA, with multiple output connectors, onto the PC system. For example, in case of our exhibitions of Media City Seoul, as described in Section 2, PC system with i7-2600 @ 3.40GHz CPU, 16GB RAM, 4TB HDD and 128GB SSD, and AMD Radeon HD 7950 3GB GDDR5 VGA (5 video output) was employed.

6.2 Software system

For the image processing and 3D visualization, we used Processing, a Java based programming tool. Max/MSP receives the fingerprint data from Processing applets via Open Sound Control [14] over UDP. Software performance is critical because Digiti Sonus must simultaneously conduct many tasks such as fingerprint sensor control, fingerprint sonification, and real-time 3D rendering. Processing has a severe limitation of being able to use only one CPU core, which we worked around by dividing our implementation into several independent applications communicating over OSC: fingerprint sensor control application, fingerprint sonification application, display application, and broadcasting application. This allows the operating system to allocate CPU core resources dynamically. An additional benefit of this architecture is that it is scalable in the number of simultaneous displays, which might change depending on the circumstances and conditions of any particular exhibition gallery.

7. EVALUATION AND RESULTS

Based on visual feature extractions and sound synthesis, we experimented sonification of fingerprints with hundreds of saved fingerprints acquired from the previous exhibition. Among those fingerprints, we selected five distinct fingerprints, which show each unique characteristic of fingerprint visual features. Each fingerprint has different ranges and numbers of parameters, and we could examine the diversity of sound in each different fingerprint. The detailed parameter numbers and 3D image results are shown at Figure 9, and it also shows the spectrograms of the sounds resulting from this same set of fingerprints.

For example, the first fingerprint resulted in the carrier frequency in low frequency range, with a harmonic set of partials because of 1.0 as the harmonicity ratio. Overall long amplitude decay made a metallic bell-like tone. The second one is similar to the first but with a higher frequency range. The third fingerprint resulted in the most surreal timbre with very low carrier frequency range and low amplitude. The fluctuating modulation index of this fingerprint resulted in vibrating sound. The fourth one was similar to the third one due to the angle of fingerprint, however the push pressure was too weak and overall amplitude was low. As a result of fifth one, a low modulation index, a short duration, and a characteristic envelope created spooky fluctuating sound. Hence the timbres of sound were all different with their own signatures, which accomplished our goal that allows users to experience distinct sonification of their fingerprints.

We had also an informal user study in Elings Hall, University of California, Santa Barbara. Digiti Sonus was installed for five hours in a hallway, and about twenty participants, who mostly have background in music, engineering, and art, experienced this installation and filled out a survey. Most participants answered that they could easily understand how their fingerprints were transformed into sound, and visuals helped to make a connection between sound and fingerprint as well. However, recognizing the comparison with other fingerprints by sound was not very easy. It was because participants mostly observed very few numbers of fingerprints, which was not enough to listen and compare dynamically different timbre of sound. Some participants who had distinctive fingerprints created surprising sound due to the unique values of fingerprint angle, brightness and number of minutiae. They were satisfied with the result and could observe the distinction with other common fingerprints sound. Some of the feedback included ideas of morphing animation when acquiring a new fingerprint, and adding text information on the screen about how minutiae are analyzed and how to make different sound with a fingerprint.

8. CONCLUSION

We have developed an interactive system generating distinct sounds for each different fingerprint using FM synthesis. It embodies fingerprint sonification and 3D visualization, including novel techniques for user-intended fingerprint expression to affect audio results. In order to enable personalized sonification and more diverse timbres, the installation employs sound synthesis based on visual feature analysis such as minutiae extraction, area, angle, and push pressure of fingerprints. Sonifying different fingerprint images results in diverse timbres, such that users could easily hear differences among fingerprints. Digiti Sonus provides each participant with a distinct individual sound based on his or her personal body data, and users can “perform” the presentation of their fingerprints as well as control the way notes are triggered in order to make the output as diverse as possible.

As future work, Digiti Sonus should be developed to give more diverse sound results. Additional sound synthesis techniques, such as granular, wave terrain, or additive, could be applied to broaden the variety of sound. Dynamic filters such as low/high pass filters should be considered as well. Future versions should also map fingerprint centroid to the visualization and/or sonification. In terms of hardware setup, the ZFM-20 fingerprint sensor reads images in low resolution and the reading speed is not fast enough. A more advanced fingerprint sensor with higher resolution and speed should be researched. It will improve the quality of visuals and sound, and enhance users' experience in better quality.

9. ACKNOWLEGMENTS

This work is the result of DaVinci Media art project 2012, which is supported by Seoul Art Space and Geumcheon art space. We thank them for all their help and support. And we thank audience who participated in DaVinci exhibition and participants who helped Elings hall informal user study for all their great feedbacks.

10. REFERENCES

[1] Han, Y. and Han, B., Digiti Sonus: an interactive fingerprint sonification. In Proceedings of the 20th ACM International Conference on Multimedia, Oct. 2012.

[2] Martoni, D. Maio, D. Jain, A.K. and Prabhakar, S. Handbook of Fingerprint Recognition. Springer, 2009.

[3] Worgan, M. Sonification of Faces. CO620 Research Project Report. Jun. 2005. Retrieved at Apr. 12, 2012 from https://www.cs.kent.ac.uk/pubs/ug/2006/ .

[4] Manabe, D. Face Visualizer. Retrieved from: http://www.daito.ws/en/work/smiles.html .

[5] Weinberg, G. and Thatcher, T. 2006. Interactive sonification of neural activity. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME06), 2006.

[6] Angel, C. R. Creating Interactive Multimedia Works with Bio-data. In Proceedings of the International Conference on New Interfaces for Musical Expressions (NIME 2011), 2011.

[7] Hong, L. Wan, Y. and Jain A. Fingerprint image enhancement: Algorithm and performance evaluation. IEEE Trans. on Pattern analysis and Machine intelligence. 20(8): 777-789, 1998. DOI=http://dx.doi.org/10.1109/34.709565.

[8] Lee, H.C., and Gaensslen R.E. Advances in Fingerprint Technology. New York, NY: Elsevier, 1991.

[9] Pearson, K. On Lines and Planes of Closest Fit to Systems of Points in Space. Philosophical Magazine, 2 (11): 559–572, 1901.

[10] Chowning, J. The synthesis of complex audio spectra by means of frequency modulation, Journal of the Audio Engineering Society, 21(7), 1973.

[11] Roads, C. The Computer Music Tutorial, MIT Press, 1995.

[12] Hermann, T., Hunt, A., and Neuhoff, J. G. The Sonification Handbook, Logos Publishing House, Berlin, 2011.

[13] Fingerprint sensor ZFM-20. Retrieved from: http://www.adafruit.com/products/751 .

[14] Wright, M. and Freed, A. Open Sound Control: A New Protocol for Communicating with Sound Synthesizers. In Proceedings of the International Computer Music Conference (ICMC), 1997.