SoniControl: Gesture Recognition System for Electric Guitar Using VLF Beacon Signals

Sangwon Suh, Juhan Nam and Sung-Hee Lee

Graduate School of Culture Technology, KAIST
335 Gwahangno, Yuseong-gu, Daejeon, Korea
[suhsw1210, juhannam, sunghee.lee]@kaist.ac.kr

ABSTRACT
As the types of sound effect units for electric guitars have been increasing, diverse methods to control them have been developed. While the most widespread interface is the pedal-type controller, this study suggests *SoniControl,* a hand motion-based effect control system. *SoniControl* is a low-cost and easy-to-build control solution for live guitar players. Using the portable media players and earphones, guitarists can easily use this new control method without physically modifying the electric guitar. The system utilizes very low frequency (VLF) range of electromagnetic (EM) wave as a beacon signal, which is generated from the media player and transmitted from the moving coil of the earphone, to calculate the position of the player's hand relative to the body of the guitar. The beacon signal is received by the guitar pickup along with its instrumental sound. The sound processor, connected to the guitar like other conventional effect chain units, recognizes gesture information from the input signal and modulates the guitar sound according to the gesture. Using this system we developed three kinds of gesture-based input for live guitar players.

Author Keywords
Musical interface, gesture, VLF, audio effect, electric guitar

ACM Classification
H.5.5 [Information Interfaces and Presentation] Sound and Music Computing – *Systems,* J.5 [Arts and Humanities] – *Performing arts*.

1. INTRODUCTION
Since audio effect units have been used for live performances, new devices have been developed to augment guitar play. A "stomp box" or "pedal" is the most common interface to control audio effects during the live performances. However, because these pedal-type effect modules are fixed on the floor, the guitarists cannot move far from them; somewhat ironically, the method to extend musical expressions constrains the movement of the guitarists. Researchers have tried to solve this problem by developing a novel type of guitars. For example, the RG Kaoss of Ibanez includes MIDI controller on the body of the guitar [2]. However, this approach cannot be applied to general guitars.

In this study, we introduce a new audio effect control system, SoniControl, which does not modify the original structure of the electric guitar. SoniControl works based on hand-motion tracking and gesture recognition using the existing pickup in an electric guitar and an earphone attached to one hand. Specifically, the moving coil of the earphone transmits very low frequency (VLF) range of electromagnetic wave as a beacon signal when a media player connected to the earphone produces a sinusoid. The system captures the beacon signal via the guitar pickup and obtains hand-motion tracking information from its strength and patterns. Since this controls audio effect by hand motions, musicians can maintain their playing style and instrumental tone without being distracted by locating controllers.

NIME'15, May 31-June 3, 2015, Louisiana State University, Baton Rouge, LA.

2. BACKGROUND
2.1 Guitar Pickup
The guitar pickup is a transducer that converts the vibration of steel strings to electric signals and transmits them to the amplifier or the analogue-digital converter. The majority of electric guitars use a magnetic pickup: a permanent magnet wrapped with a coil of several thousand turns of copper wire. The permanent magnet creates a magnetic field; the vibration of steel strings disturbs the field, and the magnetic flux induces a voltage in the coil.

The pickup is basically used to capture the sound from strings but also can receive another information. For example, TonePrint® is an effect control solution developed by TC Electronics that utilizes the guitar pickup as a near-field communication signal receiver, without physically modifying the guitar or employing add-ons [11]. With the speaker of the smartphone, musicians can transmit a preset to the connected effect modules.

2.2 Earphone
We utilize the earphone comprising a moving-coil driver: commonly called a dynamic driver. The electromagnetic force between a stationary magnet and a coil actuates the diaphragm, which then hits the air molecules to produce sound wave. In this study, we take advantage of the coil attached to the diaphragm inside the driver. Although it is unperceivable, this small part produces the electromagnetic (EM) wave to the air while the diaphragm generates the sound wave. Since the guitar pickup cannot sense the vibration of the air molecules, the EM wave from the coil is the only source that the pickup can receive. This phenomenon can be easily seen when putting the earphone to the guitar pickup; the sound from the earphone is received by the pickup and played from the guitar amp speaker. In the same way, the earphone can be used as a signal transmitter for the near-field communication with the guitar pickup.

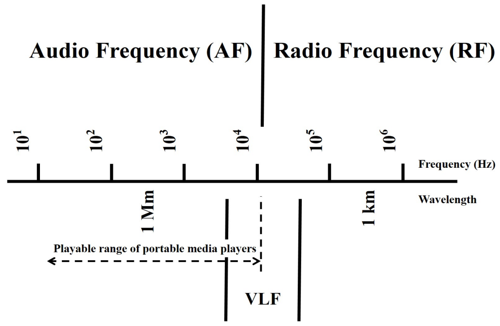

Figure 1 VLF range

2.3 Very Low Frequency (VLF)

VLF refers to the radio frequencies (RF) in the range of 3 kHz to 30 kHz: wavelengths from 10 to 100 kilometers [9]. Since there is not much bandwidth in this band of the radio spectrum, audio information cannot be transmitted, and only low data rate coded signals are used. VLF is used for communication with submarines near the surface, radio navigation beacons and time signals, and electromagnetic geophysical surveys [1].

Since the sampling rates of typical portable media players are usually 44.1 kHz, VLF band can be utilized as a near-filed communication beacon signal. Figure 1 shows the frequency range of audio frequency and radio frequency with the range of VLF. In this system, the beacon signal received by the guitar pickup passes through the beacon pass filter to extract the beacon signal from the guitar sound. If the frequency band of beacon signal belongs to the *Practically Audible Range (PAR),* then it is hard to filter the beacon signal without distorting the guitar sound. Therefore, the system utilizes *Typically Audible but Practically Inaudible Range (TAPIR)* as VLF beacon signal frequency band [12].

3. RELATED WORK

Microphones and speakers are designed to capture and generate the movement of the air molecules. In other words, if any movement interferes the sound wave traveling from a speaker to a microphone, the wave is affected by the movement. In this aspect, Gupta et al. built an air-gesture sensing system named SoundWave [10]. The system consists of a pair of built-in speaker and microphone of a laptop computer, and the speaker generates tones between 18-22 kHz, which are inaudible [7, 8]. The microphone captures the reflected signal by the movement of hand using the Doppler shift analysis. Additionally, the amplitude of the observed signal provides the information related to the proximity and the size of the target.

Musicians often make exaggerated movements to emphasize the live performance dramatically and coordinate themselves with the music. For guitarists, a soloing lead player raises his or her guitar to the near-vertical position. Willet focused on this motion and suggested a system in which the motion of an electric guitar is used to control sound effects [13]. By mapping the accelerometer data to the sound effect parameters, the sound effect parameters alter according to the three-dimensional orientation of the guitar head. The concept of tracking the orientation of the headstock as a motion controller is followed by Donovan et al [6].

Hot Hand® is a motion-based guitar sound effect control solution [4]. This system offers a ring-type wearable sensor that includes accelerometer inside. This ring device measures the movement of picking hand and transmits the data via a wireless communication module.

Sonicstrument is a motion-tracking system for musical performance using stereotypical transducers (like earphones, speakers and microphones) [5]. We initially designed our system by applying their idea to guitar effect control by utilizing the guitar pickup as a receiver. However, instead of the Doppler effect that they exploited for motion tracking, we use the strength of the beacon signal directly as this approach turned out to be more playable in our scenario. Furthermore, we develop the system to recognize pre-defined gestures and use them as control signals.

4. SYSTEM DESIGN

4.1 System Overview

The proposed system consists of an earphone as signal transmitter, a guitar pickup as signal receiver, and a PC as sound processor (Figure 2). An earphone attached to the player's picking hand plays the role of a beacon signal transmitter. The earphone should be faced with the body of guitar; the signal travels straight forward (Figure 3). The player can move the hand back and forth to control the parameter of the effect module or send commands to the system by making predefined gestures. Then, the guitar pickup on the electric guitar receives the beacon signal and transmits it to the sound processor. The input channel of the sound processor receives the signal from the guitar pickup and the output channel transmits the modulated guitar sound to the amplifier. The sound processor performs two functions: (1) detects the beacon signal and converts it to the control signal, and (2) modulates the guitar sound.

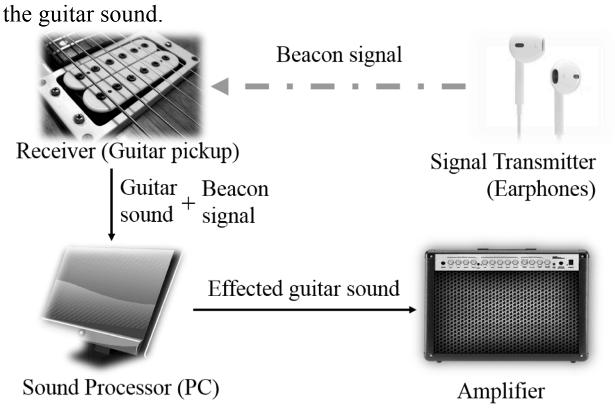

Figure 2 System overview

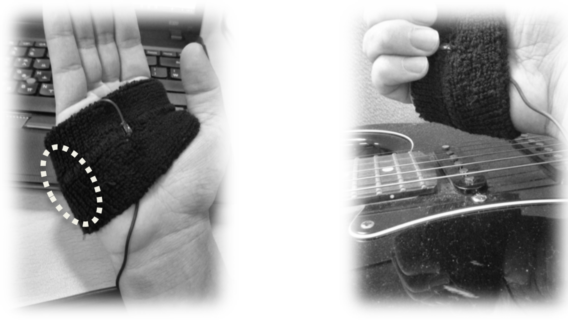

Figure 3 Installing an earphone to the player's hand. The coil of the earphone shown in the dotted circle generates the beacon signal to the guitar pickup of electric guitar (posing like a Karate Chop).

4.2 Signal Transmission

The earphone, which is connected to the portable media player, continuously generates 18 kHz sinusoidal waves to the air. At this time, two types of signals are generated; one is sound wave generated from the diaphragm and the other is EM wave from the coil. Previous work used the sound wave for sound localization and audio data transfer [3, 14]. However, we use the EM wave because the guitar pickup inherently captures sound by electromagnetic induction.

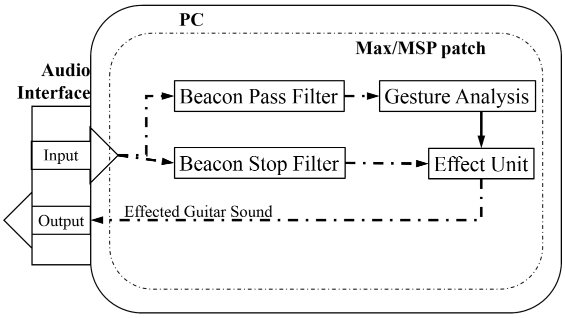

Figure 4 Block diagram of sound processor.

4.3 Sound Processor

Figure 4 shows the block diagram of the sound processor. As aforementioned, the pickup takes the mixture of the guitar sound and beacon signal. Therefore, a separation process is required. The beacon signal is extracted via the combination of high-pass and band-pass filter (18 kHz center frequency), which is termed as beacon pass filter. On the other hand, the guitar sound sources are obtained via the beacon stop filter, which eliminates the 18 kHz frequency band. Although this can reject the original guitar sound near 18 kHz, human hardly recognizes the sound in the range. The beacon signal goes into the gesture analysis module. This calculates distance information or recognizes the gesture input by analyzing the strength of the beacon signal. The analyzed control signals are sent to the effect unit to modulate the guitar sound.

4.4 Hand Tracking

We track hand motions with the distance between the guitar pickup and the earphone attached to the hand. The distance is derived from the strength of the beacon signal. Figure 5 shows the relation between the distance and the input signal strength. The diamond points represent the mean values of measured signal strength at a set of distances. Theoretically, the signal strength of spherical wave is inversely proportional to the distance from the observer. Therefore, by fitting the theoretical curve to the experimental data, we get the relation between the distance and the signal strength as follows:

$$\text{strength}(x) = -k * log\frac{x}{c} \quad \text{(A)}$$

$$k = 33.962, \qquad c = 9.0055$$

Where the input x refers to the distance and strength is in decibel. Using equation (A), the system tracks the distance from the measured signal strength and, as a result, tracks the hand motions. Figure 6 shows an example of distance tracking over time.

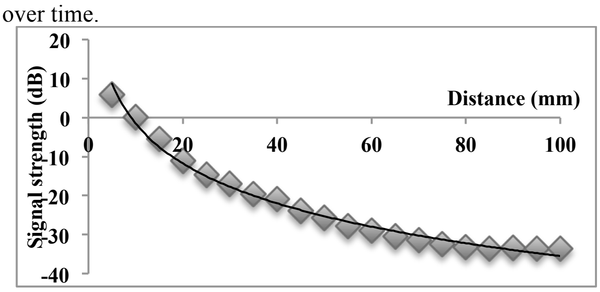

Figure 5 Strength of measured beacon signals at a set of distances and a fitted curve

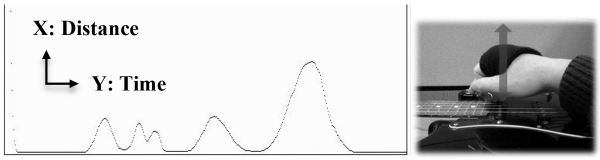

Figure 6 Tracked distances over time

4.5 Gestures

The effect unit, which modulates the guitar sound, is controlled by control signals from the gesture analysis module. The control signals include on/off command, audio effect switch command and parameter control (Figure 7).

- On/off command: This gesture sends a trigger signal to the on/off toggle switch of the effect unit. Users can input this command by bumping the face of earphone and the surface of guitar pickup. In practice, however, users will slightly push the strings by the earphone. The system recognizes the gesture by analyzing the moving average of distances between the earphone and the guitar pickup over one second. When the moving average value becomes lower than the distance between pickup and strings, the gesture analysis module triggers the on/off toggle in effect unit.
- Switch command: This command requires a two-step gesture. The beacon signal has a strong linearity. Therefore, if the earphone, located near, is not facing the guitar pickup, the system recognizes the distance as being far away. By utilizing this phenomenon, sudden change of the facing of earphone with respect to the pickup appears to large change of distance in a very short time. Users can input this command by holding the position for a second and then twisting the wrist to face the earphone to the surface of guitar pickup. By analyzing the combined data set of moving average and moving standard deviation, the system recognizes the switch command.
- Parameter control: While the effect unit is working, this gesture controls the effect parameter in real-time. For example, guitarists can control the wah-wah effect at any position using hand motions instead of stepping on the foot pedal at a fixed position.

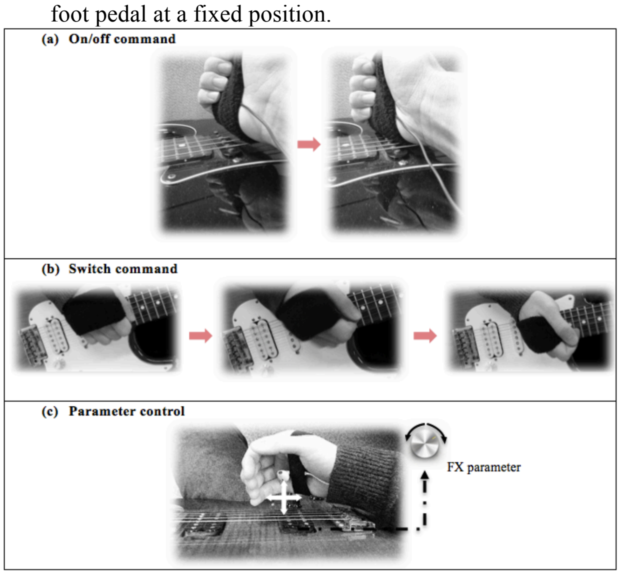

Figure 7 Gesture instructions. (a) On/off command gesture is triggered when the earphone touches the surface of the pickup (right). (b) Switch command gesture is triggered when the wrist starts to turn (middle). (c) Parameter control alters the effects parameter when the effect unit is working.

5. Switch Command Recognition

While the parameter control gesture can be used simultaneously with play motion, for example, by moving the hand horizontally while stroking, the on/off and switch commands are used between play motions; the user should stop stroking motion at least for a second before starting to input the commands. Therefore, the system should discriminate the commands from the play motions. The on/off command is easy to recognize because it is activated only when the position of the earphone is right close to the pickup for one second, which is not typical during guitar play. On the other hand, the switch command is not trivial to recognize because this command gesture occurs in the distance range of normal guitar play such as stroking. We handle this as a binary classification problem, which discriminates the switch command from the normal guitar play. To this end, we recorded the distance data during a set of test performances including switch command gestures (Figure 8). We then collected one class of examples from the switch command gestures and the other class of examples from the normal guitar play. Specifically, we extracted three features from the distance data, including moving average and moving standard deviation given a time window. Figure 9 shows the two-dimensional data plot of the features for the switch command gestures and the normal guitar play. We obtained the boundary using linear support vector machine, a popularly used simple classifier.

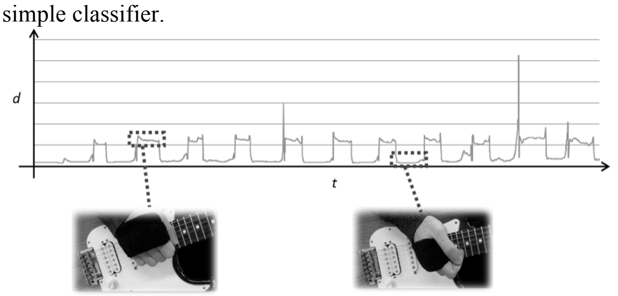

Figure 8 Time versus distance plot of n-time execution of switch command gesture. Each square represents on instances of the gesture, and the command signal is triggered when the value drops suddenly after the stationary state. X-axis is time and Y-axis is distance.

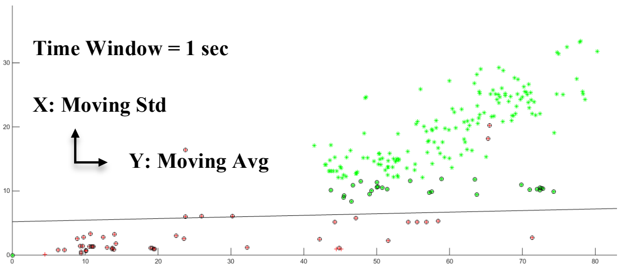

Figure 9 Moving average versus moving standard deviation graph of processed data at one second of time window. Green dots represent the stroke motion and red dots represent the switch command. The solid line represents the division line, which is determined by the support vectors (=circled points).

6. CONCLUSION

In this study, we introduced a system that captures the hand motion of guitarists with the VLF EM beacon signal transmitted from the earphone attached on the player's hand. The system utilizes the guitar pickup as a signal receiver; therefore it does not require any physical modifications or attachments on the guitar body. This feature enables almost any kind of electric guitar to become a motion-tracking interface.

The input signal of the guitar pickup is a mixture of the beacon signal and guitar sound. The system takes only the beacon signal out from the signal mixture using the beacon pass filter and returns the distance between the earphone and the guitar pickup based on the hand tracking function. In turn, the system recognizes predefined hand gestures to control the sound effect module by analyzing the distance. In case of the switch command gesture, we extract moving average and moving standard deviation for a given time window. The feature space provides a good separation between the switch command gesture and the normal guitar play.

7. REFERENCES

[1] Geonics Limited. VLF Receivers. Available at http://www.geonics.com/html/vlfsystems.html, December 23rd, 2014.

[2] Ibanez. RG Kaoss. Available at http://www.ibanez.co.jp/products/eg_page14.php?year=2014&area_id=3&cat_id=1&series_id=33&data_id=270&color=CL01, January 22nd, 2015.

[3] J. Jeon, G. Chae, E.J. Lee and W.S. Yeo. TAPIR Sound Tag: An Enhanced Sonic Communication Framework for Audience Participatory Performance. *In Proceedings of the 2014 NIME*, London, UK, 2014.

[4] J. Remignanti. *Hot Hand Motion Controlled Effects.* Available at http://www.sourceaudio.net/whitepapers/motion_control.pdf, December 25th, 2014.

[5] J.S. Lee and W.S. Yeo. Sonicstrument: A Musical Interface with Stereotypical Acoustic Transducers. *In Proceedings of the 2011 NIME*. Oslo, Norway, 2011.

[6] L.B. Donovan and A.P. McPherson. The Talking Guitar: Headstock Tracking and Mapping Strategies. *In Proceedings of the 2014 NIME,* London, UK, 2014.

[7] Lerder, C. Francis. The Human Ear. Encarta Encyclopedia Deluxe, CD-ROM, 2004.

[8] Microsoft Encarta Encyclopedia. *Ear.* CD-ROM, 2000.

[9] Rec. ITU-R V.431-7. *Nomenclature of the frequency and wavelength bands used in telecommunications.* ITU. Retrieved 20 February 2013.

[10] S. Gupta, D. Morris, S. Patel and D. Tan. Soundwave: using the Doppler effect to sense gestures. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)* (May 5-10, 2012). ACM Press, May 2012, 1911-1914.

[11] TC Electronic. TonePrint. Available at http://www.tcelectronic.com/toneprint, December 9th, 2014.

[12] W.S. Yeo, K. Kim, S. Kim and J. Lee. TAPIR Sound as a New Medium for Music. *Leonardo Music Journal 22,* 2012.

[13] W. Willett. Motion-Based Control for Guitar Effects Processing.

[14] Z. Liu, Z. Zhang, L.W. He and P. Chou. Energy-based sound source localization and gain normalization for ad hoc microphone arrays. *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on. Vol. 2, IEEE, II-*