Wiiolin: a virtual instrument using the Wii remote

Jace Miller & Tracy Hammond
Sketch Recognition Lab
Department of Computer Science & Engineering
Texas A&M University
{jacemiller, hammond}@tamu.edu

ABSTRACT
The console gaming industry is experiencing a revolution in terms of user control, and a large part to Nintendo's introduction of the Wii remote. The online open source development community has embraced the Wii remote, integrating the inexpensive technology into numerous applications. Some of the more interesting applications demonstrate how the remote hardware can be leveraged for nonstandard uses. In this paper we describe a new way of interacting with the Wii remote and sensor bar to produce music. The Wiiolin is a virtual instrument which can mimic a violin or cello. Sensor bar motion relative to the Wii remote and button presses are analyzed in real-time to generate notes. Our design is novel in that it involves the remote's infrared camera and sensor bar as an integral part of music production, allowing users to change notes by simply altering the angle of their wrist, and henceforth, bow. The Wiiolin introduces a more realistic way of instrument interaction than other attempts that rely on button presses and accelerometer data alone.

Keywords
Wii remote, virtual instrument, violin, cello, motion recognition, human computer interaction, gesture recognition.

1. INTRODUCTION
Nintendo's Wii has rejuvenated the gaming industry, due to the public's response to the controller. The Wii remote, or Wiimote as some have informally termed it, comes equipped with accelerometer sensors and an infrared (IR) camera. Leveraging this affordable technology opens up the door for new applications. This work describes a new way of interacting with the Wii remote and sensor bar to simulate playing a virtual string instrument.

This paper describes the new digital musical device, the Wiiolin. The Wiiolin is a digital stringed instrument based on the Wii technology, where the Wii remote functions as the neck, and the IR sensor bar functions as the bow. The Wiiolin automatically switches between a violin and a cello based on the orientation in which the player is holding the virtual instrument. (See Figure 1.) The angle of the player's wrist and bow defines which string the player is playing, whereas the buttons on the Wii remote define which note on the string is being played.

Figure 1: The orientation of the Wii remote determines whether the Wiiolin performs as a violin or cello.

2. RELATED WORK
A new genre of video games based on music production through instrument simulation was created with the release of the first Guitar Hero game developed by Harmonix in 2005. Since then, there have been other successful ventures in the field of video game instrument simulation, such as Rock Band and Donkey Konga. While these titles have achieved a level of popularity, they rely on instrument simulation through additional peripheral devices. Requiring consumers to purchase extra controllers in order to experience a game limits the size of the user base. Additionally, they are quite simplistic, removing many of the natural affordances that exist in playing the real instrument.

One solution to this problem is to use the standard game controller as the instrument. The most notable attempt at using the Wii remote in an instrument simulation game has been Wii Music [6]. Wii Music allows the tempo and pitch of a song to be altered using information gathered from button presses and the controller's accelerometer.

Johnny Lee's Head Tracking for Desktop Displays project uses the sensor bar in a nonstandard, "reverse" way – holding the Wiimote fixed and primarily moving the sensor bar [5]. The Wiiolin combines the ideas of instrument simulation and using the sensor bar in a nonstandard way to create a new quintephone – an electronic instrument that generates sound by optical or mechanical computing, specifically in case, using physical interaction to guide intelligent musical feedback. It provides a higher accuracy of simulation than existing systems, in that it allows users a greater amount of fluid expression and musical freedom than existing systems through the ability to change both notes and instruments through direct physical interaction with the virtual Wiiolin just as the player would with the physical device. Additionally, the Wiiolin works without additional peripherals and does not require any device other than the standard game controller and sensor bar.

3. DESIGN

A primary goal for the Wiiolin was the ability to mimic features of music production found in actual string instruments. These features include the idea of bow movement across strings to produce sound, and note pitch being determined by finger position on a string. While the Harmonix Guitar Hero removed all multi-string interaction, we chose to distinguish between multiple strings by measuring the angle of the bow compared to the virtual instrument (aka Wii remote). This is a significant improvement over the Harmonix Guitar Hero, which only allows the user to play five notes in total and using only one variance in interaction, the guitar buttons. The Wiiolin uses three buttons and four strings to allow for sixteen notes per instrument in total. The instruments belonging to the violin family, which include the violin, viola, cello, and bass, all have four strings. For each of these instruments, the standard tuning of strings form perfect fifths. Playing the open strings of a violin with no finger positions produces the notes G, D, A, and E. Standard first position, which corresponds to the key of D major, was chosen as the model for mapping finger positions to Wii remote buttons. Incidentally, the relative spatial positions of the buttons chosen for mapping finger positions are similar to actual standard first position positions. The buttons mapped to first and second finger positions are farther apart than the buttons mapped to second and third finger position. (See Figure 2.)

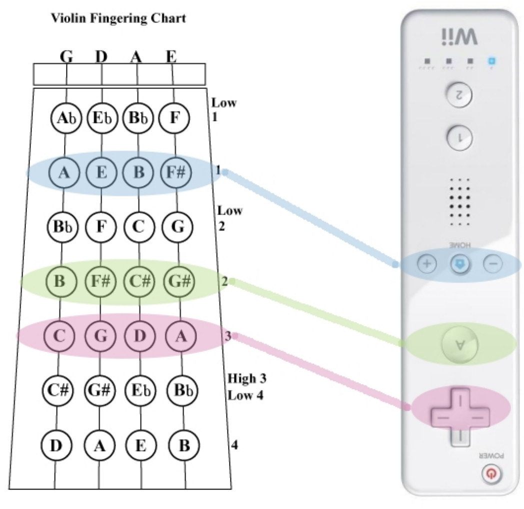

Figure 2. Mapping violin finger positions to Wii remote buttons.

Another goal was the support of multiple stringed instruments through a single device and interface. In order to automatically switch the instrument being played, we used the angle of the Wii remote, since violin players naturally hold a violin horizontally away from their body while playing it, whereas cello players naturally hold the cello vertically, resting on the floor between their knees. The result is a modeless switch between two different instruments, where players can switch instruments only by changing the position of which they hold the virtual instrument.

We also wanted to achieve realistic sound quality. While MIDI output is often used for computer generated music because it allows easy, instantaneous production of notes with the ability to sustain the note for an arbitrary length of time, we wanted to capture the sound quality reminiscent of the timbre of a string instrument. In the realm of interactive gaming, much of the processing power is devoted to rendering graphics and diverted away from sound processing. This makes processing intensive methods, such as physically modeled sound synthesis, less desirable. The goal of achieving realistic sound quality while maintaining low processing overhead guided the decision to support playback of recorded sound samples [8].

4. IMPLEMENTATION

The Wii remote is connected to a computer using a MSI Star Bluetooth USB adapter. Once connected, the Wiiolin program is executed, which continuously loops through performing three main tasks: gathering input from the Wii remote, examining the inputs to determine what note to play, and playing back the correct sound sample. Connecting to the Wii remote in a programming environment and gathering its raw input values was aided by the use of the Wiiuse C library version 0.12 [10]. The FMOD Ex 4.10 sound library was selected as the means of loading sound files into memory and handling their playback [3]. Determining what note to play consists of analyzing the state of the inputs read from the Wii remote, following the algorithm outlined below:

1. Determine the instrument
2. Determine the string
3. Determine if proper bowing motion is occurring
4. Determine the finger position
5. Select the note to be played

The type of instrument is selected based on data collected from accelerometer sensors. Values of the three accelerometer axes can be analyzed to determine the tilt of the controller with respect to the ground, given the effects of Earth's gravity. When the Wii remote is held horizontally, violin sound samples will be used; but when the Wii remote is held vertically, cello sound samples will be used. Figure 1 shows the different postures for playing the Wiiolin in its violin and cello forms. The threshold for switching instruments was set as forty-five degrees below the horizontal.

The positions of points of infrared light produced by the light emitting diodes (LEDs) in the sensor bar and detected by the Wii remote's infrared camera are used to determine the virtual string that should potentially be played. Only sources of infrared light are visible to the IR camera, which can track up to four points of IR light at a time. The sensor bar is commonly used to provide two sources of IR light, whose positions are plotted in a two-dimensional plane perpendicular to the IR camera's line of sight. The slope of a line through these points, which we refer to as the sensor bar slope, is calculated and used as the feature that selects the string. The slope space is partitioned into four sections – one for each string – with three thresholds.

The direction of proper bowing motion is dependent on the string being played. Specifically, the notion of sensor bar slope needs to be understood before bowing motion can accurately be detected. For bowing motion to produce a sound, it should occur along a vector nearly parallel to the sensor bar slope. Otherwise, moving the bow perpendicular to the normal bowing direction would also produce a sound. The equivalent action for a real world instrument would be producing sound by lifting the bow off of the strings. This erroneous situation can be avoided by finding the projection of the vector of motion onto the sensor bar slope and using its magnitude as the movement speed instead of the magnitude of the vector of movement before being projected. When the projected movement speed of the sensor bar is above a threshold, correct bowing motion is signaled.

Finger positions are determined by which buttons are pressed on the face of the Wiimote. Three buttons are mapped to both first and third finger positions to allow a greater flexibility in hand position. Just as the note produced by a real string instrument is determined by the finger position pressed nearest the bridge, the note produced by the Wiiolin is determined by the button being pressed nearest the infrared camera.

The Wii remote's infrared camera has a viewing angle of approximately forty-five degrees. This small field of view limits the range that a sensor bar can be moved while keeping both IR sources in the camera's sight. Examining a sensor bar through an IR sensitive digital camera reveals that each side of the bar actually contained three LEDs. The solution to LED spacing was to cover the center LED on one side of the bar and cover all three LEDs on the other side of the bar with the user's hand. This simple modification to the sensor bar can be seen in Figure 1. Using the closer LEDs reduced the spacing by a factor of eight, allowing a larger range of bowing motion to be detected by the IR camera while the sensor bar remains relatively close to the Wiimote.

Timing the playback of recorded music is difficult for interactive media. The authors wanted to use real sound samples in the play back rather than MIDI sounds. However, real sound samples contain short, but varied lengths of silence in the recordings before the note became audible as the bow stroke gained strength and reached the final note. We noticed that playing originally used raw sound samples from the beginning resulted in what could be perceived as a sluggish response, despite the immediate actual reaction time of the system. The time delay between the start of the sound sample and the point where it reached its recognizable pitch hampered the interactive feel of the system. In order to make the real-time response of the Wiiolin transparent to the user, adjustments to the audio samples also had to be made in the interest of simulation timing. To alleviate this problem, the first 300ms – about the average time they reach their recognizable pitch – were deleted from the sound samples. This improved the perceived response time by eliminating much of the silent portion of the recordings, but has the slight drawback of sometimes making transitions between notes sound too abrupt to be natural. A more customized solution might be to use a sound editor to remove the specific length of silence from the beginning of each sound sample. Ideally, we would have a more sophisticated bow stroke algorithm that would take into account bow speed and pressure as well as vibrato to create an even more realistic sound, as mentioned in the future work.

5. RESULTS

We performed an informal user study on our prototype Wiiolin. Twenty individuals ranging from age eighteen to fifty-two tested and tried to play the Wiiolin. Their abilities to play the Wiiolin were quite varied. All were able to produce notes when using both the violin and the cello configurations, but only those people who had experience playing a violin or cello were able to play recognizable melodies. The users found the cello position to be an easier instrument to play than the violin. We think this is due to the fact that the distance from the infrared camera to the sensor bar is generally greater when the Wiiolin is being played as a cello. All users greatly enjoyed playing with the Wiiolin and wanted to continue experimentation, enjoying the challenge of getting the Wiiolin to play actual songs.

While there is obviously a learning curve to the Wiiolin, the authors note that some of the difficulty experienced when learning bowing motion is similar to the challenge of learning the bowing motion required to play a real stringed instrument. Since the Wiiolin uses buttons to map the process of choosing what note to play to a discrete task, the learning curve is much shorter than that of an actual instrument where players must learn to memorize and precisely control continuous finger positions. The discrete mapping of notes improves note selection in a way similar to the frets on a guitar, the advantage being excellent accuracy in note pitch.

The general consensus of all the play testers was that the idea was novel and interesting. Some suggested that the prototype has the potential to be developed into a marketable product. The positive feedback has encouraged continued development of the concept, especially in regards to improving the user experience and extending the functionality of the system.

6. DISCUSSION & FUTURE WORK

This instrument is an advancement over previously existing similar virtual instruments, such as the Harmonix Guitar Hero, in that it offers the player a larger freedom of expression. Guitar Hero only has the ability to play five different notes, and thus buttons are overloaded, making it impossible for a player to play a desired song separate from their note-by-note system. In the Wiiolin users are able to play sixteen different notes through four distinct strings by the twist of the wrist and angle of the bow. This offers a more natural physical affordance than other systems.

That said, many further advancements can still be made to approve the affordance of the Wiiolin. One of the large differences that we notice between the Wiiolin and a violin is in the physical support of the instrument. A violin is held in place with nothing but the chin and shoulder. The left hand does not hold up the violin, but is used only for counteracting the force of finger presses on the strings. For the Wiiolin, the left hand must support the Wii remote, and keep it steady. Keeping the Wii remote level about its z-axis is made more important since the selected string is determined by the slope of a line in the x,y-plane. A potential solution to this problem would be to build a frame that rests on the shoulder, extends to the left hand, and attaches to the Wii remote to hold it steady. The frame could also be designed with a notch to rest the sensor bar on when bowing, yet does not obstruct the view of the IR camera.

Another feature of a real violin that we have not yet incorporated is the ability to play two strings at once. This technique is called double stopping and produces a chord of two notes. The ability to play double stopped notes could be achieved for the Wiiolin by detecting when the sensor bar slope is very near the currently defined thresholds for distinguishing strings.

A possible solution to the problem of abrupt transitions between notes may be found in tracking the acceleration of the sensor bar and predicting note changes due to changes in bowing direction. When the sensor bar accelerates in the opposite direction of bowing motion, the volume of sound samples could be faded out. The acceleration of the sensor bar in the direction of bowing motion could signal sound samples to fade in.

Other related possible improvements include simulating vibrato and using the Wii remote's internal speaker for sound playback. Vibrato is an advanced technique achieved by moving the finger quickly back and forth along the string while maintaining the pressure to produce clear notes. For the Wiiolin, this action could be cued by small back and forth acceleration along the Wiimote's z-axis. Although it is not known for great sound quality, the Wii remote's small internal speaker is supported by the FMOD Ex library and could be used to play the sound samples, removing the requirement of computer speakers.

While we would hope that this could be used to teach bowing technique while simplifying fingering technique, we question that suggestion because we expect that the absence of any physical strings may make any knowledge very difficult to transfer to a physical device. Perhaps future work that includes haptic feedback could overcome this barrier.

Currently, only two instruments are supported, but more could easily be added by incorporating sound samples from different instruments. Although any instrument could be added, the two that most naturally fit the physical mapping of the current setup are the remaining members of the violin family – the viola and the bass. The difficulty of adding more instruments is determining a way to modelessly switch to playing the additional instruments that are held similarly to the existing instruments. The violin and viola are held more or less horizontally when played, while the cello and bass are held vertically. Our solution to this is to use, in addition to Wii remote tilt, the distance from the Wii remote to the sensor bar as could be a feature used to automatically determine which instrument should be played. Calculating the distance from the Wiimote's camera to the sensor bar can be accomplished since the distance between the two infrared LEDs is fixed. The higher and lower pairs of instruments could be classified based on Wii remote alignment, while the separation between violin and viola, or cello and bass, can be done by setting a threshold for sensor bar to Wii remote distance. This threshold would most likely have to be a function of where the center of the two IR LEDs lies along the Wii remote's x-axis since the distinction of distance between the violin and viola would be fairly small. In general, the sensor bar to Wiimote distance of the viola should be farther than that of the violin. However, the sensor bar to Wiimote distance for the viola when the center of the sensor bar is closest to the Wiimote (when aligned along the Wiimote's z-axis) is less than the sensor bar to Wiimote distance for the violin when the center of the sensor bar is far from the Wiimote (when the IR LEDs are near the peripheral edge of the Wiimote camera's view). Some experimental testing will need to be done before the threshold values needed to separate the violin and viola instruments can be determined.

7. CONTRIBUTIONS

The Wiiolin provides a new way of interacting with the Wii remote and sensor bar to produce music, by changing how the user will interact with the game by providing a more intuitive physical affordance. Although other, more simplistic, virtual instruments have been developed using the Wii remote, the Wiiolin is the first to use the infrared camera and sensor bar in such a way that accurately pantomime the bowing motion and fingering to produce the desired effect on music creation. In addition, the Wiiolin allows the player to modelessly switch between different instruments. The real-time implementation lends itself to the possibility of a future interactive game being developed from the basis of the prototype.

8. ACKNOWLEDGEMENTS

This work is supported in part by NSF 0757557 and Rockwell Collins. The authors would also like to thank the members of the Sketch Recognition Lab and Kerre Cole for their help and support in the completion of this work and document.

9. REFERENCES

[1] Bencina, R., Wilde, D., and Langley, S. Gesture - Sound Experiments: Process and Mappings. In *Proceedings of the 2008 Conference on New Interfaces For Musical Expression,* Genova, Italy, 2008.

[2] Blaine, T. The convergence of alternate controllers and musical interfaces in interactive entertainment. In *Proceedings of the 2005 Conference on New Interfaces For Musical Expression,* National University of Singapore, Singapore, 2005, 27-33.

[3] FMOD Music and Sound Effects System. [Web site] 2010, http://www.fmod.org.

[4] Paine, G. Towards unified design guidelines for new interfaces for musical expression, In *Organised Sound 14,* 2 (Aug. 2009), 142-155.

[5] Lee, J. C. Head Tracking for Desktop VR Displays using the Wii Remote. [Web site] 2008, http://www.cs.cmu.edu/~johnny/projects/wii/.

[6] Nintendo Official Website - Wii Music. [Web site] 2010, http://www.wiimusic.com/launch/index.html.

[7] Peng, L. and Gerhard, D. A Wii-based gestural interface for computer-based conducting systems. In *Proceedings of the 2009 Conference on New Interfaces For Musical Expression,* Pittsburgh, PA, USA, 2009.

[8] Philharmonia Orchestra Sound Sample Library. [Web site] 2010, http://www.philharmonia.co.uk/thesoundexchange/sound_samples/sample_libraries/.

[9] Toyoda, S. Sensillum: an improvisational approach to composition. In *Proceedings of the 2007 Conference on New Interfaces For Musical Expression,* New York, NY, USA, 2007.

[10] Wiiuse – The Wiimote C Library. [Web site] 2008, http://wiiuse.net/.

[11] Young, D. The Hyperbow controller: real-time dynamics measurement of violin performance. In *Proceedings of the 2002 Conference on New Interfaces For Musical Expression,* National University of Singapore, Singapore, 2002, 1-6.