Exploring 3D Gestural Interfaces for Music Creation in Video Games

Jared N. Bott
University of Central Florida
School of EECS
Orlando, Florida 32816 USA
jbott@cs.ucf.edu

James G. Crowley
University of Central Florida
School of EECS
Orlando, Florida 32816 USA
jcrowley@cs.ucf.edu

Joseph J. LaViola Jr.
University of Central Florida
School of EECS
Orlando, Florida 32816 USA
jjl@eecs.ucf.edu

ABSTRACT

In recent years the popularity of music and rhythm-based games has experienced tremendous growth. However almost all of these games require custom hardware to be used as input devices, and these devices control only one or two similar instruments. In this paper we describe One Man Band, a prototype video game for musical expression that uses novel 3D spatial interaction techniques using accelerometer-based motion controllers. One Man Band provides users with 3D gestural interfaces to control both the timing and sound of the music played, with both single and collaborative player modes. We further investigate the ability to detect different musical gestures without explicit selection of mode, giving the user the ability to seamlessly transition between instrument types with a single input device. A formal user study is then presented comparing the musical interface of One Man Band to that of Nintendo's Wii Music. Our results indicate that users generally preferred the interface of One Man Band over that of Wii Music. We also found that users desire to express their own ideas and have explicit control of the melodies created in music-based video games.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Auditory (non-speech) feedback*; K.8.0 [**Personal Computing**]: General—*Games*; I.5.5 [**Pattern Recognition**]: Implementation—*Interactive Systems*

General Terms

3D gestural interfaces, video games, music, Wii Remote, user evaluation

1. INTRODUCTION

Recently in the gaming industry there has been growing interest in music and rhythm games. Franchises such as Guitar Hero and Rock Band have sold over a million units, with newly released titles continuing to generate significant

interest. These games offer the ability to simulate playing music by using custom input devices designed to emulate the instrument being played. Although it has been shown that such props can lead to a more immersive experience [7], these input devices represent an additional investment required of the player. Furthermore while these games may give the player the feeling of actually playing an instrument, there is almost no opportunity for individual musical expression. Instead, the user is relegated to using the input device to trigger a predefined sequence of notes. Thus, most games of the genre are more about matching rhythms than actually creating music.

Figure 1: **Four people playing One Man Band in guided mode.**

We posit that it is possible to create a music game that gives the user more creative control while preserving those characteristics that have made rhythm games popular. To explore this idea we have created One Man Band, a prototype video game that uses 3D gestural interfaces to simulate a number of different musical instruments. Our prototype uses Nintendo Wii controllers, that utilize accelerometers, to identify gestures made by the players and translate them into musical notes or sounds. An important design goal in One Man Band was to enable musical novices to collaborate and create pleasing compositions.

Our prototype contains several different gameplay options, including a guided mode, a free play mode, and a mechanism for a single player to create complex songs on their own. In guided mode, players can choose a song and an instrument and are presented with graphical cues indicating which notes to play. Free play mode allows players to

explore their own musical ideas and offers a visual representation of any notes played. A looping mechanism is provided to each player so that they can create complex rhythms and songs on their own or as part of an ensemble. Finally, One Man Band provides a special gestural interface that allows a single player to use different instruments without explicit selection of mode.

We evaluate the One Man Band interface by presenting both a short informal study of the effectiveness of playing instruments without explicit selection of mode, and a formal user study comparing the One Man Band interface to that of Wii Music, a commercially available game from Nintendo that also uses a Wii Remote as an input device. We conducted our experiment by having subjects play instruments from both interfaces and gathering feedback about what aspects of the interfaces they liked and disliked. Subjects ranked their preferences on a seven point Likert scale and we used standard statistical techniques to analyze the data.

2. RELATED WORK

Interactive music games have been developed and studied by both the academic community and the entertainment industry. Guitar Hero [1] and Rock Band [2] are just two recent examples of music-based games that have experienced tremendous success. Both of these games provide the user with specialized input devices designed to mimic the instrument they represent. Players are asked to press buttons on a guitar controller or strike drum pads to match symbols that appear on screen. If the player performs the correct action with proper timing the in-game music continues to play; if the player instead fails to press the correct button or does so with incorrect timing, a non-musical sound is triggered instead. Guitar Hero has also recently introduced a studio mode that allows players to develop their own musical ideas using the supplied input devices. For an excellent discussion of additional games and interfaces in this space see [4]. Although each of these games provide some collaborative and creative abilities, our work also gives the player a simple and powerful gestural interface for many instruments using only a single input device.

Wii Music [3] is a recent title that focuses on using a simple interface to allow musical expression. This software uses an accelerometer-equipped input device to recognize musical gestures. However, Wii Music is limited with respect to melodic expression. Players are not able to compose music. Instead, players choose from a list of songs, and use gestures to define rhythms that are applied to notes selected by the software itself. This is a significant omission, as users are provided no mechanism to actually *create* music. Our work also uses simple gestures to implement musical interfaces, but additionally gives users creative control over exactly what melodies and rhythms they want to play.

The topic of interactive music has also been explored by researchers. World Beat [6] provides players with infrared batons that can be used to conduct a virtual orchestra or control several instruments. Researchers in [13] also examine digital batons as gestural interfaces. CoJIVE [8] explores musical collaboration by allowing novices to use digital batons to collaborate on jazz compositions. Jam-O-Drum [5] also focuses on the collaborative aspects of music making, placing users at a tabletop installation where they can participate in a number of interactive music games by striking
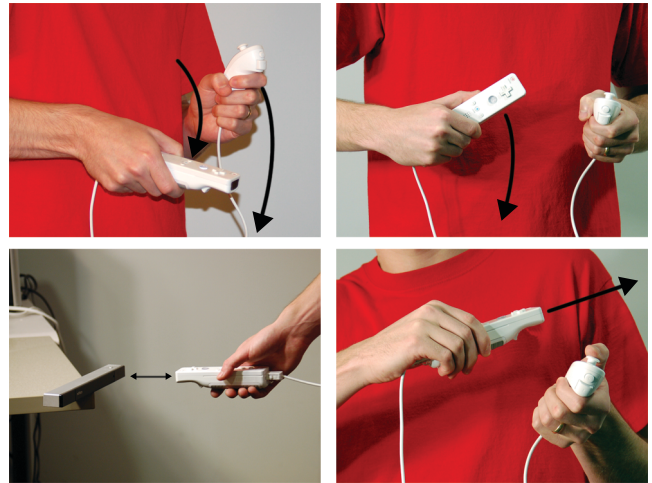

Figure 2: Musical instruments are played through several gestural interfaces. Clockwise from top-left: drums, guitar, violin, theremin

a drum pad. Additional projects have also explored the concept of interactive or collaborative musical systems [9, 12, 14]. Our work builds upon these ideas by using off the shelf hardware and the PC platform to implement a collaborative gesture-based music game that does not require custom interfaces.

3. IMPLEMENTATION

One Man Band uses a game controller in the form of Nintendo Wii remotes and nunchuks, a common accessory for the Wii remote. Wii remotes contain a three-axis digital accelerometer, an infrared camera, a directional pad, and a number of buttons. The nunchuk, which plugs into the Wii remote, contains the same accelerometer as well as a thumbstick and two buttons.

Our prototype provides six primary single-instrument gestural interfaces: drums, guitar, violin, bass, trombone, and theremin. These interfaces allow players to produce sounds by mimicking the way each instrument is played (Figure 2). The interfaces are designed to give the user an intuitive feel for how each instrument type should be used. This scheme allows for an interface that is easy, realistic, powerful, and expressive. We carefully chose the notes and sounds of the instruments so players can effortlessly play and combine sounds together in a way that is both musical and pleasing to the ear, and we provide several sets of sounds for each instrument type so that the player can create songs in a variety of genres.

3.1 Interfaces

Drums. The drums are implemented as virtual drumsticks, with a gesture simulating a drumming motion along the Z-axis of the device triggering a drum sound. At first we simply played a drum sound whenever we saw a sufficiently large acceleration peak, but this allowed players to trigger drum sounds with an upward motion as well as a downward motion. Furthermore, anytime the Wii remote changes direction or stops, two distinct peaks will be seen, one with a positive acceleration and another with a negative acceleration.

To solve both of these problems we implemented a state machine that tracks the motion of the device. A drum sound is only triggered when the Wii remote is moving down with an acceleration greater than 1.5 g. The nunchuk is also used as a virtual drumstick in the same manner, giving the user the ability to generate interesting rhythms by using two input devices in tandem.

Guitar. This interface is implemented much like the drums but with a different orientation. The Wii remote is held with the buttons facing outward, and a strumming motion is used to generate chords. The particular chord is specified by orienting the nunchuk in one of three positions and pressing zero, one, or two buttons. Although we had originally wanted to allow users to play via both down and up strums, we discovered that simply returning to a still position would often trigger an unintended chord. We therefore only allow chords to be triggered with strums in the downward direction, and we use the same state machine described above to detect the direction.

Violin and Bass. This pair of instruments are both played by using the Wii remote to make a bowing gesture along the Y-axis. We once again implement a state machine to track the current motion, but we allow both forward and backward strokes to be played. This is more feasible for this particular interface because the direction of motion is orthogonal to the gravitational force reported by the accelerometers. Therefore, this stream is more precise than that of instruments that operate with an up and down motion. This allows us to use a smaller threshold (only 0.4 g) while still differentiating between a backward bow and a return to the original orientation. The violin and bass play sounds from the same key as the other instruments, with the note specified by orienting the thumbstick on the nunchuk in one of eight positions. The octave of the note can be changed by pressing two buttons, giving the user a total of three octaves of the given scale.

Trombone and Theremin. Trombone and theremin share similar interfaces in that they both operate by pointing the IR camera on the Wii remote at an infrared light source. The camera's distance from the light is used to determine the pitch of the sound played. The trombone instrument maps seven notes into a portion of the total distance range and the portion can be changed through a calibration step. The player presses a button to trigger the chosen sound, and another button toggles the octave of the note played. The theremin interface plays a continuous tone and uses the entire distance range to control pitch. Hearing this tone at a constant volume is less than desirable, so we give the user the ability to expressively control the volume through the orientation of the nunchuk, with the percentage of full volume heard calculated as the square of the Z component of its accelerometer, saturating at a value of one.

3.2 Gameplay

As mentioned previously One Man Band provides two primary gameplay modes, a guided mode and a freeplay mode. In both guided mode and freeplay mode, a metronome is used to help players keep time. The metronome can be configured by changing the time signature or tempo, and can also be disabled if desired. In guided mode (Figure 1), we provide several songs of varying difficulty for players to choose from. Each player also chooses from one of four instruments available for each song; no two players can play
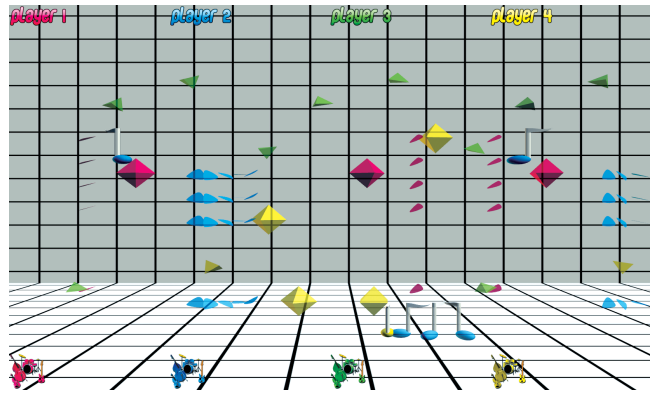

Figure 3: Notes are represented graphically in freeplay mode.

the same part, although some songs have two parts for one instrument. The metronome is used to provide visual feedback in the form of scrolling lines representing beats in the song. A thick blue line at the top of the screen denotes the current time. Suggestions as to what might be played take the form of symbols that scroll on the prescribed beat. These icons give players a guide to playing the song, like sheet music; however, players are free to improvise as much as they would like and any melodies played will be heard regardless of what notes are suggested.

Freeplay mode allows up to four players to improvise using the provided musical interfaces. Visual cues can often aid inexperienced musicians in understanding what they and others are playing [5], so we provide players with a three-dimensional scene containing notes and geometric shapes representing the tones and rhythms played. When a player creates a note an object is added to the world and that object travels through the three-dimensional space (Figure 3). Each instrument has its own object type and each type has its own unique patterns of motion determined by the speed of the metronome. For instance, some models bounce as they move through the world, landing on each beat. Notes with a low pitch appear toward the bottom of the scene, while notes with higher pitch are placed near the top. These objects help players to choose notes that interact with the other players' visual output. Players can switch between two views, one with objects moving across the screen from left to right, and another where the objects move away from the player, into the screen, and eventually disappear into the distance.

3.3 Looping

In freeplay mode, each player is also given the ability to create repeated patterns of notes. Any sound the user plays may optionally be recorded and looped at an interval chosen by the user. This interval can be as small as a single measure, or arbitrarily large. Once the player has recorded one or more notes she can choose to quantize the recorded notes to the nearest beat or any subdivision of beats. By applying quantization the player can build a rhythmically steady background track over which to improvise. We have found this ability to be particularly suited to novices who may have trouble playing accurate rhythms. By recording a melody or drum line and applying quantization even the most inex-
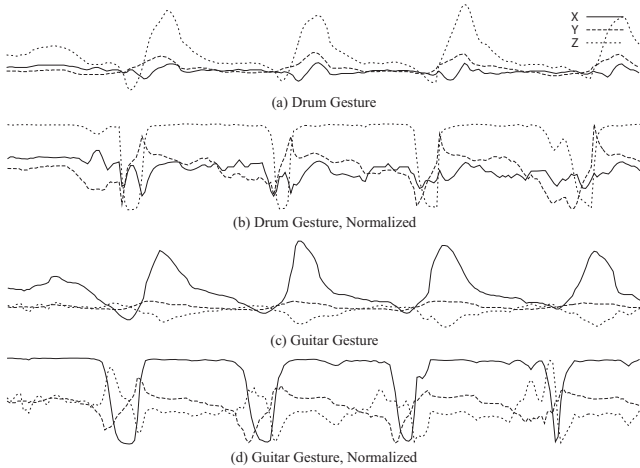
Figure 4: Raw and normalized accelerometer data for both the drum and guitar interfaces. The data shows four distinct drum gestures in (a) and (b) and four distinct guitar gestures in (c) and (d).

perienced musician can develop rhythmically pleasing songs, giving the user confidence to explore their ideas further.

All functionality required to create loops is implemented through the Wii remote. The user can enable and disable recording or playback, undo any number of recorded notes, or clear notes by instrument. This seemingly simple mechanism gives the players a powerful method to create musical patterns and explore complex compositions, and these musical creations can be saved for playback or editing at a later time.

3.4 MIMI: a Multi-Instrument Musical Interface

All of the previously discussed gestural interfaces are chosen through explicit selection. However, we wanted to remove this limitation and provide seamless transitions between instruments while still using a single input device. To accomplish this, we needed the ability to detect when a user plays a particular instrument. Although there are many known techniques for gesture classification [10], the problem of differentiating between musical gestures in real time is slightly different from general classification.

MIMI has an additional constraint that the gesture must be recognized at a precise moment (a peak in the accelerometer data), and this may be difficult to accomplish using traditional gesture recognition techniques. Furthermore, we speculated that simpler techniques might suffice for the relatively small number of gestures we wished to classify. Therefore, we decided to investigate a heuristic-based approach to recognize these musical gestures.

We first tried to distinguish between only two gestures: a drumming motion, and a guitar strumming motion. These gestures are very similar except for the axis of motion. That is, they differ only in orientation. Keeping track of the last known orientation is simple, but requires the user to pause in the correct orientation since orientation can only be precisely determined when an accelerometer-based input device is stationary. This prevents the user from moving fluidly from one instrument to another and thus was not considered a suitable solution. Instead, we needed to develop a

method of tracking both approximate orientation and motion concurrently.

An instrument will generate a play event when its state and the current data trigger such an event, as described in section 3.1. However due to non-precise gestures and accelerometer jitter it is common to trigger a play event for one instrument while intending to play another. The problem therefore is one of classification; we allow all instruments to be active simultaneously but only allow an instrument's play event to generate a sound if the classifier reports that this instrument is being played.

Our classification function is implemented by passing accelerometer data through a series of mathematical functions. The magnitude of acceleration is not important to the classifier, as we are only trying to track the dominant axis of motion, so we first remove the dramatic peaks in the stream of accelerometer data by normalizing each accelerometer vector such that:

$$\vec{a}_{norm} = \frac{\vec{a}}{\parallel \vec{a} \parallel}. \tag{1}$$

Figure 4 shows a sample of accelerometer data for drums and guitar, both before and after normalization. The raw accelerometer data clearly shows the timing of each musical gesture, while the normalized data gives an indication of the dominant axis of movement and orientation. Once the data is normalized we pass it through an exponential smoothing function

$$\vec{a}_{current} = \alpha\vec{a}_i + (1 - \alpha)\vec{a}_{i-1}, \tag{2}$$

using an α of 0.1 to remove jitter as in [11]. Note that there is an inherent tradeoff between removing as much noise as possible and timely classification. If the stream is not smooth enough classification will be hampered by the noise present. Conversely, if too much smoothing is applied the classifier will take an unacceptable amount of time to generate the correct classification when moving between instrument types.

This type of classification works well for gestures that use a downward motion, as these gestures have similar output both in a still state and when in motion. However the violin gesture we have implemented has an orientation that is orthogonal to its axis of movement. This leads to frequent misclassification when using the technique above. To solve this problem, we tried using smoothing functions with different parameters. We discovered that an α of 0.5 allows accurate classification of short, sharp gestures that may be missed using a lower α. This is especially true for our violin gesture, since the orientation of the device does not differentiate it from drums. Using these two smoothing functions together, the first to roughly determine orientation and the second to capture faster movements made with ambiguous orientations, we were able to accurately differentiate between the three types of motion.

To support a fourth instrument we leverage the fact that the thumbstick on the nunchuk is used when playing bass. We overload the guitar strumming gesture such that when the thumbstick is pressed it is interpreted as a bass; otherwise it is a guitar. Finally, we implement a fifth instrument type by banging the Wii remote and nunchuk together. The sound produced differs with each instrument set. For instance, in the orchestral sound set a hand cymbal is heard;

a more dance-oriented set instead produces a hand clap.

4. PRELIMINARY EVALUATIONS

During development of the prototype we had several people, with a variety of musical backgrounds, try out our software both in the guided mode and freeplay. Participants were generally able to use the provided musical gestures with little or no instruction. We found that people tend to intuitively use the Wii remote to imitate the instrument that they are playing, although some required coaching, especially with respect to proper orientation of the device for each instrument. For instance, many users tried to play the guitar with the device facing upward, while we designed this interface to be used with the top of the Wii remote facing outward. Subjects also expressed their general satisfaction about the ability to create pleasing rhythms and melodies, despite their perceived lack of musicality.

Participants tended to express more excitement when playing the guided mode than when playing in freeplay mode. This is somewhat contrary to our hypothesis that users desire the ability to express themselves creatively within a music game. However we theorize that this preference is simply a manifestation of the desire to have a tangible goal; we suspect that if users were given an opportunity to spend an extended period of time with the prototype they would feel a stronger desire to explore freeplay mode. Those who did spend some time in freeplay were especially fond of the three dimensional visualization.

We also asked 10 participants to take part in a small informal study to evaluate the accuracy of the MIMI instrument. In developing this interface we noticed that the classifier most frequently made a misclassification when switching from one instrument type to another. Therefore we designed an experiment to determine the effectiveness of the classifier when moving between instrument types. Participants were given the opportunity to become comfortable with using the five instruments and then were asked to play three random orderings of all five instruments, such that each instrument was played exactly once in each of the three sets.

With 150 transitions between instruments, MIMI correctly recognized the participants' gestures 136 times, or 91%, with three subjects playing their sequence perfectly. We found the largest source of error to be the variety of ways individuals interpret a gesture. For instance, when asked to move the Wii remote downward to strum the guitar some participants still moved the device with a more radial motion. Although this type of gesture might be successful some of the time, it did not yield consistent results in our study. Some individuals were also confused about when to use the thumbstick when asked to play guitar or bass, but we believe this would be resolved with more practice.

The other frequent source of error was users who were anxious to play the next instrument type and did not return to a correct orientation before making the prescribed gesture. Although the MIMI instrument does not require the user to pause in a certain orientation before playing, it does rely on the fact that users will tend to hold the Wii remote in a specified position at the beginning of a gesture. Moving from a guitar orientation directly into a drum strike, for instance, is difficult for our classifier to handle because it is expecting to see the entire strike along the same axis.

5. EXPERIMENTAL STUDY

To explore the effectiveness and usability of our musical interface, we conducted a study comparing One Man Band to Wii Music for the Nintendo Wii. The interface used in Wii Music is similar to One Man Band in that they both use the Wii Remote to provide the user a method of playing music. However the interfaces differ in both the gestures used, and how the user controls what music they create. In particular, Wii Music does not provide the user the ability to select the pitch of the notes they play.

5.1 Subjects and Apparatus

We recruited sixteen students (12 males, 4 females) from the University of Central Florida to participate in our study. The ages of our subjects ranged from 18-28, and nine of these subjects had previous experience playing a musical instrument. Eleven subjects had previous experience playing music-based video games, but none had previously played Wii Music. All but one participant had prior experience using a Wii Remote. The experiment took approximately 40 minutes to complete and each participant was paid 10 dollars for their time.

Our experimental setup consisted of a 50 inch DLP monitor used to display the software for the two interfaces. Wii Music was run on a Nintendo Wii console, while One Man Band was run on a desktop PC with a 2.4 GHz quad-core processor running Windows Vista. For both interfaces, a Wii Remote and nunchuk were used as input devices.

5.2 Experimental Task

Subjects were asked to try the violin and guitar instruments for One Man Band and Wii Music. The One Man Band interface operates as described in Section 3.1. The violin instrument in Wii Music is operated by pressing a button on the nunchuk attachment to create a note. The button can be held indefinitely to allow the note to continue playing. The volume of the note is determined by how quickly the user moves the Wii Remote using a bowing gesture.

The guitar in Wii Music is played by holding the Wii Remote with the buttons facing upward and strumming. Pressing one button on the nunchuk switches from playing a single note to a chord, and pressing the other button on the nunchuk makes the note shorter, with a plucked sound. For both violin and guitar, the pitch of the note the user hears is determined by the software itself.

After trying each interface, subjects were asked to complete a post-technique-questionnaire about their experience. Subjects rated their experience in eleven categories using a seven point Likert scale for each interface (where 1 indicated strong agreement and 7 indicated strong disagreement), and were asked to comment on what they liked and disliked. Each subject was also given a post-questionnaire with seven statements comparing Wii Music with One Man Band.

Finally, after these questionnaires were completed, each subject was asked to try out MIMI for approximately three minutes. Once they had experimented with this interface we asked each subject to comment on what they liked and disliked. We also asked each subject if they thought MIMI would be useful in a video game setting and to comment on that interface.

5.3 Experimental Design and Procedure

We used a two-way within-subjects factorial design, where

Guitar				
Question	**Z Value**	**P Value**	**Mean**	
			OMB	**WM**
Realism	-2.501	p<0.05	3.250	4.375
Behavior	-2.036	p<0.05	2.438	3.375
Expressiveness	-2.252	p<0.05	2.438	4.125
Explore	-2.562	p<0.05	2.250	4.125
Intent	-1.960	p<0.05	2.125	3.500
Notes	-2.612	p<0.05	2.313	4.313
Musical	-3.072	p<0.05	2.438	4.625
Difficulty	-1.996	p<0.05	5.938	4.625
Frustration	-2.276	p<0.05	6.313	4.750
Create	-3.093	p<0.05	3.125	5.063
Enjoy	-3.311	p<0.05	1.813	4.313

Violin				
Question	**Z Value**	**P Value**	**Mean**	
			OMB	**WM**
Realism	-2.350	p<0.05	3.000	4.123
Behavior	-1.488	p=0.137	2.813	3.750
Expressiveness	-1.867	p=0.062	2.438	3.438
Explore	-2.358	p<0.05	2.188	3.500
Intent	-0.836	p=0.403	3.467	3.250
Notes	-0.946	p=0.344	3.067	4.000
Musical	-0.759	p=0.448	3.133	3.688
Difficulty	-0.791	p=0.429	4.800	5.375
Frustration	-0.728	p=0.467	5.867	5.438
Create	-0.632	p=0.528	3.600	4.375
Enjoy	-1.361	p=0.173	2.563	3.500

Table 1: Data Analysis for Guitar and Violin.

the independent variables were the musical interfaces. The two interfaces studied were One Man Band and Wii Music. The dependent variable was user preference, which was based on a post-questionnaire that asked subjects to rate their agreement with eleven statements on a seven point Likert scale for each instrument and interface.

Initially subjects were given a pre-questionnaire. They were then asked to use the guitar and violin instruments for each interface. Half of the subjects tried Wii Music first, while the other half tried One Man Band first. Subjects were given approximately three minutes to experiment with each interface and then asked to play "Twinkle Twinkle Little Star". In both One Man Band and Wii Music, subjects played along with three prerecorded instruments. When playing Wii Music subjects did not receive any on screen instruction and were told to play whenever and however they liked. Since notes must be explicitly selected in One Man Band, subjects used guided mode (Section 3.2) and were visually shown the notes used to play "Twinkle Twinkle Little Star". Subjects were reminded that they were free to play whenever and however they liked, and the notes were simply suggestions.

5.4 Usability Metrics

Data for the experiment was primarily gathered in post-technique-questionnaires about each instrument type. The following questions were asked for each instrument and interface:

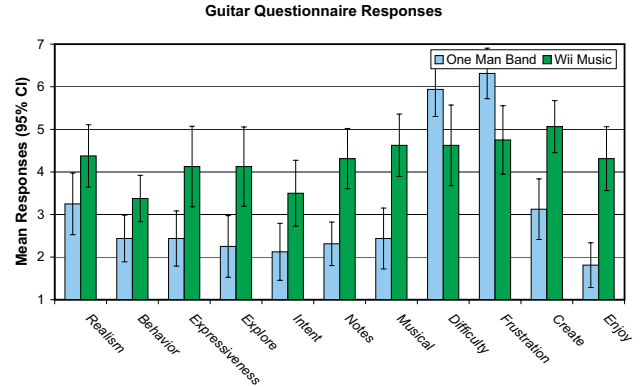- *Realism*: "I felt like I was really playing the instrument."

Figure 5: Results for the guitar instrument. A response of 1 indicates strong agreement, while 7 indicates strong disagreement. For a full list of statements evaluated see Section 5.4.

- *Behavior*: "The instrument behaved like I expected it to."
- *Expressiveness*: "I was able to easily express my own musical ideas."
- *Explore*: "I felt free to explore my own musical ideas."
- *Intent*: "I was able to play a note at the time I intended."
- *Notes*: "I was able to play the notes that I desired."
- *Musical*: "I felt my playing to be musical."
- *Difficulty*: "I had a difficult time playing Twinkle Twinkle Little Star."
- *Frustration*: "I was frustrated while trying to play Twinkle Twinkle Little Star."
- *Create*: "When I played Twinkle Twinkle Little Star, I felt like I was creating music."
- *Enjoy*: "I enjoyed playing this instrument."

Each subject was also given a post-questionnaire comparing One Man Band to Wii Music. The questionnaire contained the following:

- *Guitar*: "Which did you prefer for guitar?"
- *Violin*: "Which did you prefer for violin?"
- *Frust*: "Which did you find more frustrating?"
- *Enjoy*: "Which did you find more enjoyable?"
- *Easy*: "Which did you find easier?"
- *Music*: "Which allowed you to best express your musical ideas?"
- *Prefer*: "Which did you prefer overall?"

5.5 Results

To analyze the data collected about each interface, we performed an analysis with pairwise comparisons using Wilcoxon Signed Rank tests on the post-technique-questionnaire data. In particular, we made a comparison for both instrument types, violin and guitar, between Wii Music and One Man Band. We then performed a Chi-Square test on the post-questionnaire data that directly compares the two interfaces to determine significance.

For the violin instrument, we found only two statements to be significant. For the statement "I felt like I was really playing the violin", subjects preferred the One Man Band interface. Subjects also preferred One Man Band when presented the statement "I felt free to explore my own musical

ideas". The complete set of results is shown in Table 1. For guitar we found the responses to all eleven statements to be significant, with the subjects always favoring One Man Band. The results for the guitar post-questionnaire and a list of questions are presented in Figure 5 and Table 1.

When asked what they liked about the violin instrument from Wii Music, subjects frequently cited the fact that it was "easy" or had a "simple interface". When asked what they disliked, subjects responded that it was "boring", and that the interface was "awkward", or that "it felt like I was pressing a button, not playing". For One Man Band's violin instrument, participants most frequently identified the bowing motion as the thing they liked, with statements such as "...the bow movements were similar to real life playing" and "it gave you a pretty good feeling of playing the instrument". When asked what they disliked, participants often said that they found selecting a note with the thumbstick to be difficult.

For the guitar instrument, subjects stated that they liked how easy the Wii Music version was to use, and also that they liked the controls, stating it was "responsive" and "life-like". Subjects did not like the simplicity of the interface, saying it was "too easy" or "restrictive through simplicity". Subjects also cited the orientation of the Wii Remote (buttons facing upward) as something they disliked. When asked about the One Man Band guitar, participants liked the fact that they had control over the chords being played. The two things subjects disliked where the fact that they could not play with an up-stroke, and that they could not play a single note, only chords.

Participants were also asked to directly compare Wii Music and One Man Band in seven categories, and those results are shown in Figure 6. Performing a Chi-Square test on the data reveals a statistically significant preference for only three questions: which interface they preferred for guitar ($\chi_1^2 = 6.25, p < 0.05$), which interface they found more enjoyable ($\chi_1^2 = 6.25, p < 0.05$), and which interface they preferred overall ($\chi_1^2 = 4.00, p < 0.05$). For each of these questions participants preferred One Man Band.

Finally, subjects were asked about the MIMI interface. A statistically significant number of participants, 12 out of 16 ($\chi_1^2 = 4.00, p < 0.05$), thought the MIMI interface would be useful in a video game setting. When asked what they liked about MIMI, subjects most frequently stated that it allowed them to be creative. Subjects also found it to be "cool" and "innovative". On the other hand many participants stated that they disliked the accuracy of instrument recognition, stating that they heard one instrument when they intended to play another. Some subjects also said they found it confusing to keep track of how to play each instrument.

6. DISCUSSION

The results of our experiment indicate that users preferred the One Man Band guitar interface over that of Wii Music. The gestures used for guitar in each interface were very similar, so this would suggest that participants favored the fact that they were in control of what notes they were playing in One Man Band, as opposed to just the timing of the notes in Wii Music. Comments about the two guitar interfaces support this idea. It is interesting to note also that while many users were pleased with the creative control they had with One Man Band, a few participants disliked that they were limited to only twelve chords, while Wii Music has

many more possible sounds. This highlights the fact that although we have identified trends in preferences, it remains true that different users have very different expectations and desires about what they want in a game.
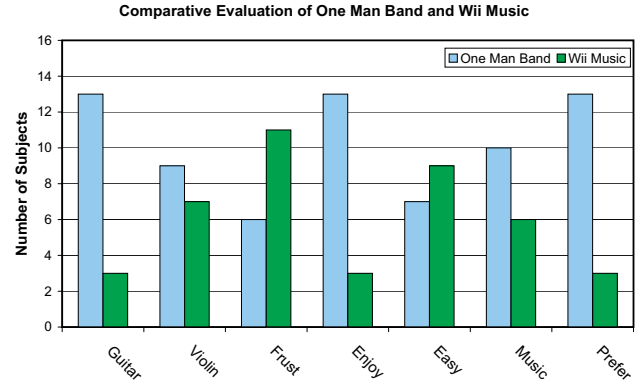

Figure 6: Post-questionnaire comparing One Man Band and Wii Music. For a full list of questions, see Section 5.4.

The results for the violin instrument were less clear, but a preference for One Man Band was observed in two categories. The lack of more statistically significant preferences for violin likely arise from the fact that the two instruments were fundamentally different in control. Coupled with the differences in ability to select the pitch of notes, there may have been too many variables present to isolate a clear preference between the two interfaces. From observations and the comments made on post-questionnaires it seems that participants tended to prefer the creative aspect of the One Man Band violin, but some disliked the controls. On the other hand many subjects liked the controls of the Wii Music violin, but disliked the fact that they could not explicitly select the note they wanted to play.

Overall we believe the results of this study demonstrate that players desire the ability to have creative control in music based video games. Subjects showed a clear preference for the guitar instrument where they were allowed to choose their own notes despite the fact that the controls for each guitar were similar. It is interesting to note that, although subjects said that they felt more creativity using One Man Band, few of these subjects actually improvised while playing "Twinkle Twinkle Little Star" and instead were content to play exactly the notes shown to them. This may indicate that what the subjects really liked was the feeling of being in control, rather than the ability to be creative. In any case, our study indicates that users prefer an interface where they have a consistent way to control the timing and the pitch of musical notes.

7. CONCLUSION

We have presented One Man Band, a prototype video game that uses 3D gestural interaction for simulating different musical instruments. Each gestural interface closely mimics a specific instrument in function and in sound, with gesture recognition performed using heuristic-based classification on Wii remote accelerometer data. We have also

described a multi-instrument musical interface (MIMI) that operates without explicit mode selection that uses a combination of acceleration normalization and exponential smoothing to determine which instrument is played.

Although we have managed to differentiate between five instruments using a single device, we plan to investigate whether even more gestures can be added to the MIMI instrument. More accurate recognition might be possible by selecting different or additional heuristics, alternate exponential smoothing functions, or more formal gesture classification techniques. A comparison between our method and current classification techniques with respect to accuracy may also be enlightening.

We have also presented a study that compares the One Man Band interface to that of Wii Music. Our results indicate that users prefer the One Man Band interface when playing in guided mode. We observed that users generally preferred the ability to explicitly specify the notes they produced. The success of titles such as Guitar Hero and Rock Band seem to contradict this finding, as they remain quite popular despite not allowing the user to explicitly control the music they play. However, we would suggest that these games instead give the user the *feeling* of having control through immersive gameplay. The perception of the user seems to be the most important factor, and if a game purports to give the user the ability to create music, our findings suggest that users desire control over both timing and pitch. We would like to investigate this idea in more detail in the future with a study that directly assesses how creativity affects a user's experience.

We believe that One Man Band shows that a simple input device can be used effectively to give the player the ability to create musical compositions. These compositions can be complex while still providing a simple interface to the user. We propose that such interfaces might be adopted in future interactive music applications to give users simple but powerful methods to express themselves musically.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] Guitar hero, February 2009. `http://www.guitarhero.com`.

[2] Rock band, February 2009. `http://www.rockband.com`.

[3] Wii music, February 2009. `http://www.wiimusic.com`.

[4] T. Blaine. The convergence of alternate controllers and musical interfaces in interactive entertainment. In *NIME '05: Proceedings of the 2005 conference on New interfaces for musical expression*, pages 27–33, Singapore, Singapore, 2004. National University of Singapore.

[5] T. Blaine and T. Perkis. The jam-o-drum interactive music system: a study in interaction design. In *DIS '00: Proceedings of the 3rd conference on Designing interactive systems*, pages 165–173, New York, NY, USA, 2000. ACM.

[6] J. O. Borchers. Worldbeat: designing a baton-based interface for an interactive music exhibit. In *CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 131–138, New York, NY, USA, 1997. ACM.

[7] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 2004.

[8] J. Buchholz, E. Lee, J. Klein, and J. Borchers. coJIVE: A system to support collaborative jazz improvisation. Technical Report AIB-2007-04, RWTH Aachen, Mar. 2007.

[9] M. Desainte-Catherine, G. Kurtag, S. Marchand, C. Semal, and P. Hanna. Playing with sounds as playing video games. *Comput. Entertain.*, 2(2):16–38, 2004.

[10] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley and Sons, 2001.

[11] J. J. LaViola. An experiment comparing double exponential smoothing and kalman filter-based predictive tracking algorithms. In *VR '03: Proceedings of the IEEE Virtual Reality 2003*, pages 283–284, Washington, DC, USA, 2003. IEEE Computer Society.

[12] R. W. Lindeman. Drumar: An augmented reality drumming game, September 2008. `http://web.cs.wpi.edu/~gogo/hive/DrumAR/`.

[13] T. Marrin. Possibilities for the digital baton as a general-purpose gestural interface. In *CHI '97: CHI '97 extended abstracts on Human factors in computing systems*, pages 311–312, New York, NY, USA, 1997. ACM.

[14] B. Vigoda and D. Merrill. Jamioki-purejoy: a game engine and instrument for electronically-mediated musical improvisation. In *NIME '07: Proceedings of the 7th international conference on New interfaces for musical expression*, pages 321–326, New York, NY, USA, 2007. ACM.