An Investigative High-Level Design of an Electric Bass Tutoring System Integrating Game Elements

Adam S. Jenkins

Music Technology Area, Department of Music Research
Schulich School of Music
McGill University
Montreal, Quebec, Canada

February 2011

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Arts in Music Technology.

Abstract

There have been a number of computer assisted musical instrument tutoring systems developed over the years for a variety of instruments. However, there has not been significant development in a tutoring system designed specifically for the electric bass guitar. Furthermore, while many of these tutoring systems leverage technology, the vast majority have not capitalized on increasing student motivation and engagement through integrating digital game-based learning.

This thesis presents a high-level design of an electric bass tutoring system which integrates game elements. In order to establish a well-founded design, the thesis first reviews: the notable design features from previous computer assisted musical instrument tutoring systems, the musical elements of an electric bass performance which a human teacher would focus on when providing feedback to a student, methods of extracting these musical elements, and the game elements which foster student engagement and motivation.

Sommaire

Au cours des dernières années, plusieurs systèmes d'enseignement assisté par ordinateur ont été développés pour une variété d'instruments de musiques. Cependant, il n'y a toujours pas de développements considérables d'un système éducatif qui soit dédié à l'apprentissage de la basse électrique. Une grande partie des systèmes existant misent beaucoup sur la technologie sans chercher à motiver l'élève en intégrant l'élément de participation interactive que peut générer un jeu électronique éducatif.

Ce mémoire présente la conception d'un système d'enseignement de haut niveau qui examine les éléments du jeu dans le cadre de l'apprentissage de la basse électrique. Afin d'établir les fondements de la conception, le mémoire considère : les principales caractéristiques des systèmes d'enseignement musical assisté par ordinateur antérieurs; les éléments musicaux d'une performance de basse électrique sur lesquels un enseignant se concentrerait au cours du processus de rétroaction en personne; les méthodes qui permettent de déterminer ces éléments musicaux, et; les éléments du jeu qui favorisent la participation et la motivation de l'étudiant.

Acknowledgments

I would like to thank my advisor, Gary Scavone, for your trust, guidance, and allowing me to work on a thesis subject that I would grow to love. Thank you to my parents Alan and Beverley, my sister Tara, my brother-in-law Branko, and nephew Connor for your continued love and support. Thank you Philippe Depalle and Marcelo Wanderley for your suggestions during my research. Thank you to my Music Tech and McGill colleagues Darryl, Mark, Federico, Bertrand, Amie, Jason, Joanne, Jung Suk, Nate, Avrum, Joe, Harry, Andrew, Dave, Jessica, and Marielle for your help and good times throughout this journey. Thank you Mikael and Jordan for your warm welcoming hospitality and friendship. Thank you Sarah and Mikael for your lovely assistance translating the abstract to French. Thank you to all my Montreal and Toronto friends who have encouraged me and shown me love throughout this arduous yet rewarding experience.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

El Sistema, founded by Jos Antonio Abreu, has shown how a unique educational experience can nurture and shape a student's love of playing music (Hollinger 2007). I have been a tutor and teacher's assistant, and have taught an undergraduate class. I have a deep interest in education and also in how technology can be used to improve education. I have been playing electric bass for fifteen years in several rock, jazz, funk, and electronic bands, and I also play piano and guitar. However I am most knowledgable and proficient with the electric bass. I am also interested in video games and I have been playing them consistently since I was a kid. With these three interests of education, music, and video games, I wondered how the use of video games could foster a learning environment that engages students while teaching them effectively how to play the electric bass guitar.

The video game industry is a multi-billion dollar industry. In recent years there has been an increased advocacy of the legitimacy of using video games as a medium for delivering educational content. Games have the ability to immerse players in a world where they can receive immediate feedback through a customized learning experience.

There are computer assisted musical instrument tuition (CAMIT) systems for piano, wind instruments (such as trumpet, clarinet and flute), stringed instruments (such as violin, viola and cello), and even guitar. However, from my research there are little that focus on the electric bass. Furthermore, many current CAMIT systems use technologies to deliver music tuition, however they do not incorporate the depth of available engagement or motivational capabilities that can be found by incorporating digital game-based learning principles. One reason these systems do not exist may be the lack of high-level design documents which outline how the two can be fused together.

An analogy can be found in the design of a sailboat. A hull is used to keep the ship afloat. A mast and sail are used to capture energy from the wind and propel the ship forward. The keel is used to keep the boat from tipping over and lastly the rudder and tiller are used to alter the course of the sailboat. Each element is necessary for the sailboat to function as it does.

Similarly, there have been studies on computer assisted musical instrument tutoring systems (which combine music technology with music education) and there has been research on combining education with digital games. However there has been little research involving the combination of music technology, music education, and digital game design, and even less regarding application to the electric bass guitar.

1.2 Goals

My goal is to present a high-level design of an electric bass tutoring system which incorporates game elements.

In order to justify the design choices, I will outline notable features of past computer assisted musical instrument tuition systems, specify a list of musical elements that a tuition system should extract from an electric bass performance, examine methods and technologies for extracting the musical elements, and examine how digital game-based learning can be used to increase engagement and learning potential.

I will not be implementing the proposed tutoring system as it would require too much development time and would be beyond the scope of the thesis. Furthermore, I will not specify the way the design must be implemented. Instead, my the goal is to specify the high-level structure of the design.

My goal is also not to introduce new technologies or methodologies. Instead, it is to consolidate current technologies and methodologies, and ultimately render a new design for electric bass tutoring system.

1.3 Thesis Overview

This thesis will include:

- A thorough survey of related computer assisted musical instrument tuition systems and their notable features (See Chap. 2)

- A description of the musical elements relevant to teaching electric bass (See Chap. 3)
- An overview of musical element extraction methods and technologies (See Chap. 4)
- An overview of engaging game elements used in educational digital games (See Chap. 5)
- A detailed proposed for a high-level design of an electric bass tutoring game (See Chap. 6)

Chapter 2

Survey of Previous Music Tutoring Systems

There have been a variety of computer-based technologies used as teaching tools to help students improve their musical skills. Brandão (Brandão, Wiggins, and Pain 1999) present a review of some computer software used in music education. Newer technologies and methodologies continue to emerge through the evolution of computer based musical education systems. When determining how various technologies can be used in the design of a computer assisted musical instrument tutoring (CAMIT) system for electric bass guitar, it is relevant to outline what notable technologies and methodologies currently exist.

The surveyed systems in this chapter were designed for instruments other than the electric bass. However, the theories, methods, and technologies in these systems can easily be incorporated into the design of the proposed electric bass tuition system described in Ch. 6.

2.1 Piano Tutor

Piano Tutor (Dannenberg et al. 1990, 1993) was a computer-based system developed to assist beginner students learning how to play the piano. The system was designed to act as a tutor which would work in conjunction with, and not replace, traditional piano teaching. It incorporated score following capabilities, an expert system, a lesson database, a design based on instructional design theory, and multimedia technologies (see Fig. 2.1). The multimedia technologies were comprised of a MIDI keyboard, music synthesizer, prerecorded speech stored on the computer, and a videodisc player. Since musical information was comprised of MIDI messages there was no need for any digital signal processing.

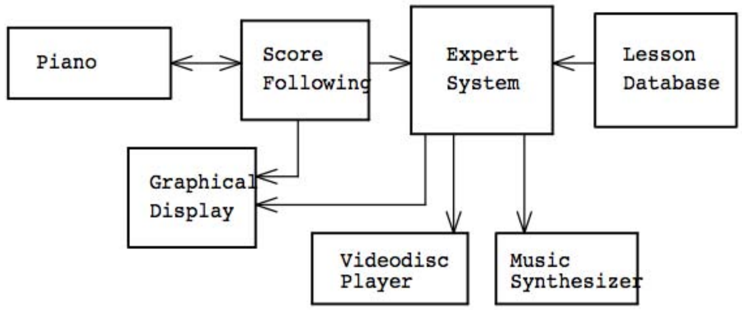

Fig. 2.1 Piano Tutor's principle components from (Dannenberg et al. 1990)

Piano Tutor's main focus is its expert system. It delivers a lesson by giving instructions to the student, evaluating student performance, offering feedback regarding the student's performance, and suggesting any necessary remedial work.

The system has two control structures: a practice mode and a teaching mode. In the practice mode, students can choose which lessons they want to work on. In the teaching mode, the system acts as a teacher suggesting lessons. A good human teacher knows what students should work on based on how well they have performed in the past. The expert system keeps track of students' skills and capabilities in a *student model*. The student model contains information about students' skills based on how well they performed historically in previous exercises. From the student model, Piano Tutor is able to suggest the most appropriate lesson to help students improve.

The structure of the lesson is shown in Fig. 2.2. The system reviews the student model and selects the most suitable lesson for a particular student from the lesson database. The software delivers the lesson with the aid of multimedia sources including videos played via a videodisc, graphics such as a musical score displayed on the computer monitor, and audio such as accompaniments or prerecorded voice recordings.

After the student is shown the content of the lesson, he or she is asked to play the music from a musical score shown graphically on the computer display. During the performance, the software performs score following and any necessary automated page turning. At the end of the exercise, the software analyzes the student's performance, prioritizes errors, and suggests any remedial exercises. A diagram of the evaluation and remediation selection process is shown in Fig. 2.3.

Errors are conveyed to the student through text comments, speech, video presentations

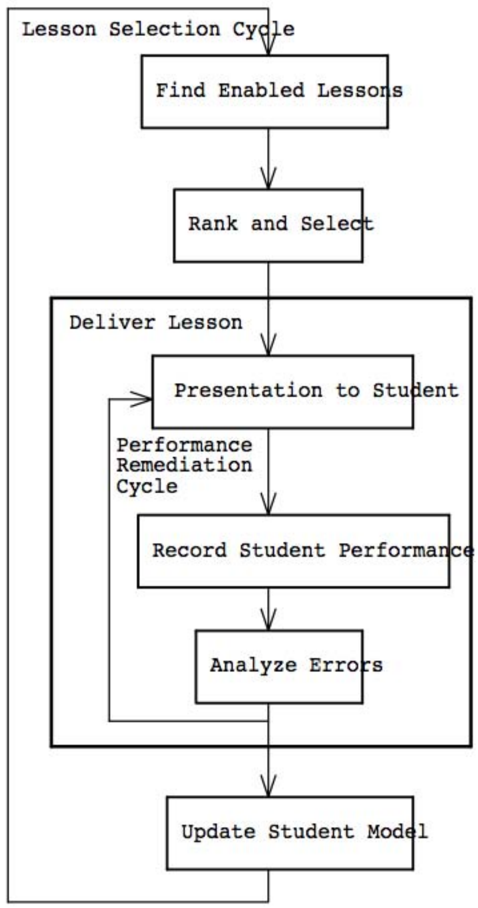

Fig. 2.2 Piano Tutor's basic control structure from (Dannenberg et al. 1990)

and/or highlighted notes on the score.

If the system detects that the student has difficulties performing an exercise, the software may request that it be done again. If the difficulties persist, the software will help the student by recommending a remedial lesson.

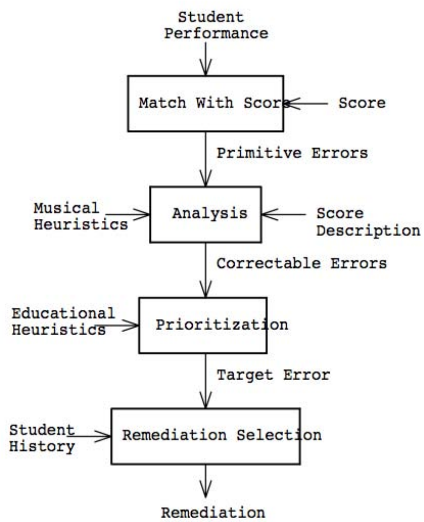

Fig. 2.3 Piano Tutor's performance evaluation and remediation selection from (Dannenberg et al. 1990)

2.2 PianoFORTE

PianoFORTE (Smoliar, Waterworth, and Kellock 1995) was another type of multimedia computer technology which allowed students to receive feedback regarding their performance of a piece of music. Similar to Piano Tutor, it used a MIDI keyboard to transmit the musical information of students' performances. Unlike, Piano Tutor, PianoFORTE did not include the function of recommending lessons. Instead, it focused on ways of effectively explaining the details regarding how the student executed and interpreted the piece.

Music is more than simply playing the correct notes at the correct time. PianoFORTE's design recognizes this and puts emphasis not simply on reading and executing music notation, but on learning how to perform music. Reading music notation is a subgoal of music performance, once students have a satisfactory literacy of music notation and are able to know *what* to play, the piano teacher then provides more information on *how* to play it.

In their paper, the authors emphasize that:

> "...as far as music education is concerned, what most needs to be acquired [by the student] is the skill of *listening*. The teacher can only assess the success of communication with the student upon hearing how the student is actually performing, and the student's own knowledge ultimately depends on an ability to hear a performance the same way the teacher does. A teacher who tries to correct a performance is usually trying to discuss certain features which the student does not yet hear. Improvement will only come when the student finally masters the ability to hear those same features." (Smoliar, Waterworth, and Kellock 1995)

With this concept in mind, pianoFORTE attempts to help students *listen* to their performances. Since sound is transitory and only exists while it is being played, pianoFORTE enables the students to record and play back their performances for closer analysis. Moreover, the system has the ability to isolate and play back the performances based on specific sections of the musical score, as opposed to selecting based on minutes and seconds into the performance. This way students' performances can be more readily compared to that of teachers' or to the students' previous performances. In contrast, if the performances were recorded using a tape recorder, digital recorder, or other forms of recording device more time would be spent trying to find specific sections. Having the ability to immediately listen back to a specific section will help maximize students' practice times.

Practice time and comprehension are also increased when details regarding the execution of performances are communicated clearly. PianoFORTE utilizes graphical representations; A typical session starts with either students recording their performances of a specific notated piece of music or with a demonstration performance. After the performance is loaded, details regarding the performance are displayed. This is useful for the students because not only does the system display whether students played the correct note at the correct time, it also displays data regarding the dynamics, tempo, articulation, and synchronization of the right and left hands.

2.2.1 Dynamics

The loudness of a note is estimated using velocity MIDI data. The velocity is denoted as *loud* or *soft* depending on whether it is above or below the average velocity. Though perceived loudness is a non-linear function of frequency, the designers of pianoFORTE chose not to take this into account. They also admitted that using velocity MIDI data would be a very rough approximation of loudness. However, for the purpose of visualizing the dynamics of the performance it was claimed to be suitable enough.

A screen shot of dynamics evaluation is shown in Fig. 2.4. The loudness is represented visually by the colour of the note head. Soft notes are shaded blue, while loud notes are shaded red. The degree to which the note is soft or loud is represented by the saturation of either blue or red. Therefore, the colouring from soft to average to loud is blue, grey, and red respectively. Notes which were not played have black note heads while a green X indicates a note was played that is not in the score.

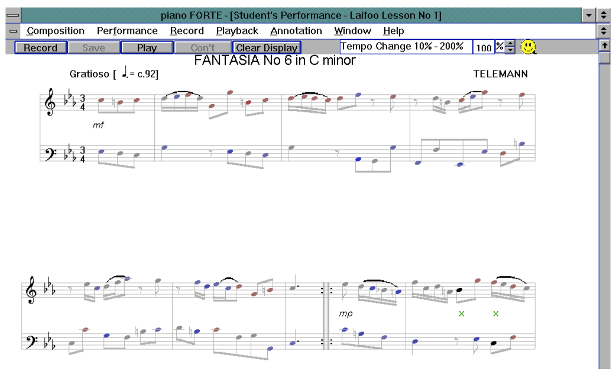

Fig. 2.4 Dynamics shown in PianoFORTE from (Smoliar, Waterworth, and Kellock 1995)

2.2.2 Tempo

The tempo of a performance is calculated by comparing the time between the onset of notes. The intervals are normalized depending on the expected note value marked in the musical score. PianoFORTE uses a *speedometer graph* to annotate the tempo curve of the performance (see Fig. 2.5). The horizontal axis on the speedometer graph is aligned with the notes on the musical score. The vertical axis is described as representing "the speed values as the inverses of the normalized duration intervals, and it is scaled to be linear in these inverse duration values" (Smoliar, Waterworth, and Kellock 1995). These speed values give an indication of the timing, however the values do not explicitly depict whether students are playing before, on, or after the beat.

The tempo values of the right and left hands are shown separately. The values for the right hand are coloured purple while the colours for the left hand are coloured orange.

A horizontal dotted green line represents the expected tempo if one is marked in the

Fig. 2.5 Tempo shown in PianoFORTE from (Smoliar, Waterworth, and Kellock 1995)

score. A faster or slower performance will be above or below this dotted line respectively. For example, in Fig. 2.5 the student is performing slower than the expected 92 dotted quarter notes per minute (as labeled in the score).

2.2.3 Articulation

The articulation of a note represents the duration time from the onset to the offset of a note (e.g. legato staccato). This duration is indicated graphically with isosceles triangles (see Fig. 2.6). The base of the triangle is aligned vertically with the centre of the note head. The tip of the triangle ends in relation to the offset of the note. Similar to the tempo markings, the right and left hand are distinguished with purple and orange colour coding. The duration of the articulation markings are strictly from when a piano key is pressed and released, irrespective of the damper pedal. Pedal markings are shown with a black horizontal line starting and ending with a black triangle.

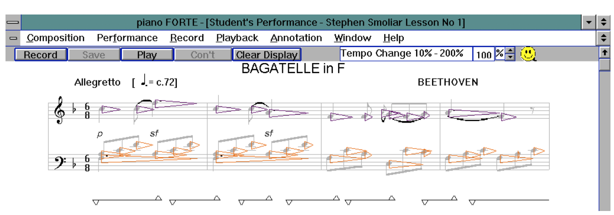

Fig. 2.6 Articulation shown in PianoFORTE from (Smoliar, Waterworth, and Kellock 1995)

2.2.4 Synchronization

The synchronization delineates whether the onset of notes played by the right and left hands occur at the same time, if the notes should be played at the same time. This is marked by a vertical line. Again, the marking for the right hand is purple while the left hand is marked as orange. The synchronicity of the onsets is shown by how close the vertical lines align with each other (see Fig. 2.7).

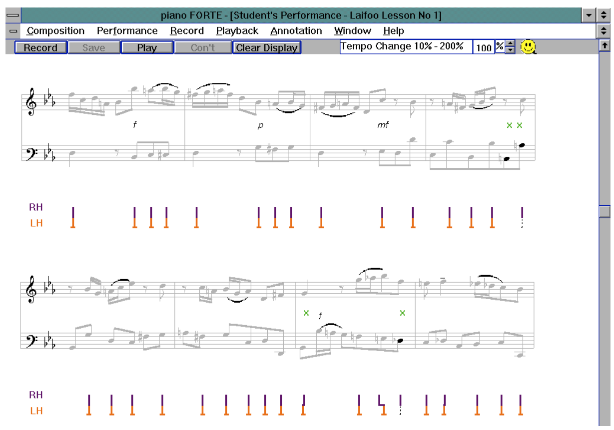

Fig. 2.7 Synchronization shown in PianoFORTE from (Smoliar, Waterworth, and Kellock 1995)

2.3 IMUTUS

The European project, IMUTUS (Interactive Music Tuition System), (Fober, Letz, and Orlarey 2005; Fober et al. 2004; Schoonderwaldt, Askenfelt, and Hansen 2005; Schoonderwaldt, Hansen, and Askenfelt 2004) aimed to create a software that could act as a tutor for a monophonic non-MIDI instrument. The recorder was selected as the non-MIDI instrument to be used for this project because it was widely taught throughout European schools. This software was not intended to replace human music teachers, but rather to work in conjunction with them as a supplementary tool to help students practice their instruments.

Similar to Piano Tutor and PianoFORTE, the students' main method of interaction with the system is a musical score (See Fig. 2.8 for a screen capture of the IMUTUS User Interface). New music scores to be used by the system could be created by teachers using IMUTUS's score editor (whose format uses MusicXML), or through a standalone optical music recognition tool. The optical music recognition tool allows teachers to scan a musical score which is transcribed into a digital format (MusicXML) to be used by IMUTUS.

Fig. 2.8 IMUTUS's user interface from (Raptis et al. 2005)

2.3.1 Three phases of interaction

Once a score is chosen, students use the *Score Viewer* to proceed through three phases of interaction: an *explorative phase*, *practicing phase*, and *feedback phase*.

During the *explorative phase* students read the score to be performed and listen to any reference recordings. The *practicing phase* which consists of performing the score, allows students to choose whether or not to use the built in metronome. During the performances, real-time audio analysis facilitates score following and automated page turning or scrolling. Finally, during the *feedback phase* the students' performances are recorded while the software performs its real-time score following and automated page turning or scrolling. The recording is then analyzed and compared to the reference score off-line for score mapping and performance evaluation. The performance evaluation module (PEM) detects performance errors, assigns an overall grading of the performance between 1 and 5, and determines appropriate feedback to be displayed on the score viewer (see Fig. 2.9).

2.3.2 Feedback

In order not to overwhelm students, specific feedback that balance both positive and negative constructive criticism. The PEM prioritizes and selects feedback based on the fol-

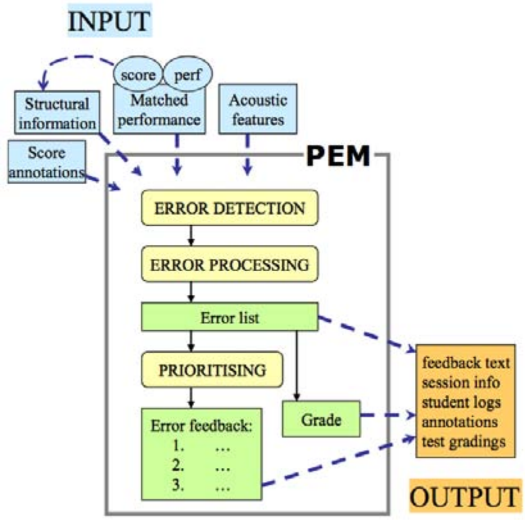

Fig. 2.9 IMUTUS's Performance Evaluation Module (PEM) design from (Schoonderwaldt, Askenfelt, and Hansen 2005)

lowing error types (listed from highest to lowest priority): air flow, fingering, rhythmic performance, attack, melodic performance, tempo, intonation, phrasing, and articulation. The order of priority is based on the results of questionnaires collected from 40 recorder teachers as well as interviews with five recorder teachers during the initial phase of the project (Fober et al. 2004). The questionnaires asked the music teachers to rate errors based on their relative importance during the first few terms of playing the recorder. The results are shown in Fig. 2.10.

No	Basic error/mistake	Average ranking	Instrument control (IC) Musical Performance (MP)
1	Air flow	1.7	IC
2	Fingering	1.7	IC
3	Rhythmic perf.	2.0	MP
4	Attack	2.0	IC
5	Melodic perf.	4.2	MP
6	Tempo	5.0	MP
7	Intonation	5.3	IC
8	Phrasing	6.0	MP
9	Articulation	7.5	IC/MP

Fig. 2.10 Basic error categories by priority from (Fober et al. 2004)

Once the three types of feedback are selected, a feedback message is selected for each from a database of standard feedback messages. If the issue is related to fingering, the system also has a virtual-reality fingering viewer (see Fig. 2.11). This allows students to have a visual representation of the fingering of specific notes.

Fig. 2.11 Fingering Viewer from IMUTUS's Dec 2004 newsletter (IMUTUS 2010)

2.3.3 Learning settings

IMUTUS was designed for three different learning settings: self-learning, classroom, or distance learning. In self-learning settings, students use the IMUTUS system to work independently on musical tasks offered by the system. This may improve students' effectiveness and motivation for practicing. In the classroom setting, the teacher can use the IMUTUS system as a teaching tool to convey musical concepts and offer feedback. In the distance learning setting, students can access and download new content from their home computers. The new content is retrieved from a content repository stored on an accessible server. The content consists of education units (theory units, exercise units, game units, and tests) created by the teachers using IMUTUS's content authoring module. Students are also able to upload their performances to the server, which can then be viewed and graded by the teachers. Finally, teachers are able to communicate with students about assigned content and view students' progress.

2.4 VEMUS

The VEMUS (Virtual European Music School) (Fober, Letz, and Orlarey 2007; Tambouratzis et al. 2008; VEMUS 2010) was a collaborative project among six countries:

Estonia, France, Greece, Lithuania, Romania, and Sweden. It built upon the platform, tools, and results of its predecessor, the IMUTUS project.

2.4.1 Chosen Instruments

While IMUTUS was designed exclusively for the recorder, VEMUS extended the number of non-MIDI monophonic instruments to include flute, saxophone, clarinet, and trumpet. These instruments were chosen based on a survey that determined which wind instruments were most widely taught in European schools (Tambouratzis et al. 2008). VEMUS added a 3D fingering viewer for each of the four new instruments (see Fig. 2.12).

Fig. 2.12 Fingering Viewer samples from (VEMUS 2009)

2.4.2 Feedback

VEMUS also extended the ways of communicating feedback to the students. Annotation capabilities were extended from graphic images, highlighted regions and text to include emoticons[1], hand written annotations[2], audio annotations[3], and real-time audio processing[4] (see Fig. 2.13).

The musical features of students' performances are also represented visually through a variety of individual or combined graphic curves. These curves attempt to help convey information regarding music elements such as articulation, nuance, pitch, timbre, intonation, stability of pitch, stability of volume, and timing. Table. 2.1 shows the correlation between these musical elements and the available graphic curves. The curves are displayed

[1] In Fig.2.13, the emoticons are the different coloured smiley face icons

[2] Hand written annotations are not shown in Fig.2.13, however they are the equivalent to drawing on the score as one would using a painting software

[3] In Fig. 2.13, audio annotations are shown by the white speaker icon with the blue centre. They allow the teacher to add audio examples for the student to listen to, which can be useful when helping the student hear how a certain passage should sound.

[4] The *freeverb* and *echo* markings shown in Fig. 2.13 are in fact real-time audio effects which are triggered when students reach those points in the score.

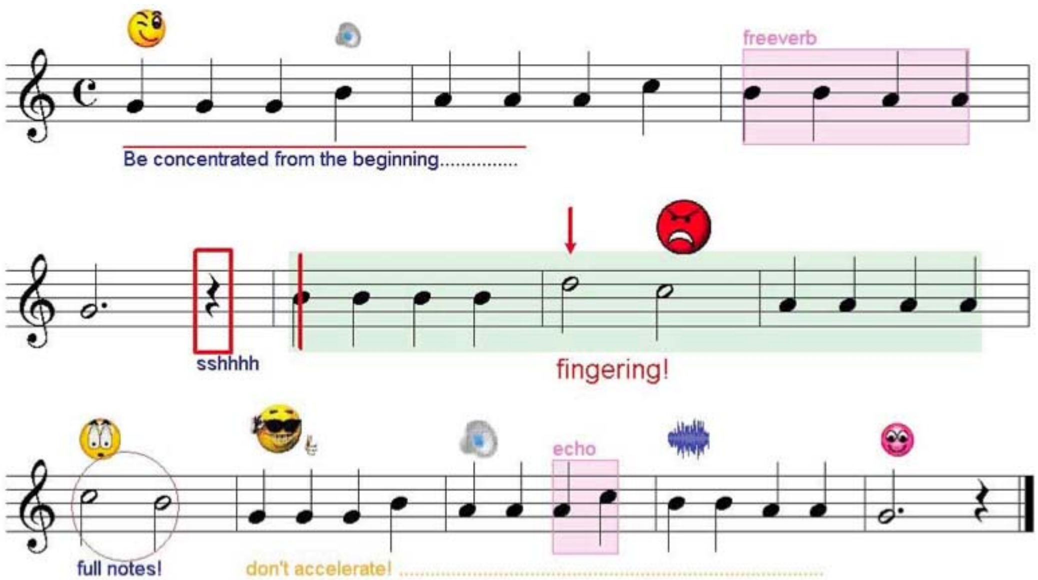

Fig. 2.13 Annotated score in VEMUS from (Fober, Letz, and Orlarey 2007)

adjacent to the musical score, enabling students to have a clearer understanding of how they performed during specific sections. An example of a combined curve which graphically displays the pitch, envelope, and spectral content of a student's performance is shown in Fig. 2.14.

curve						stability		
symbolic name	articulation	nuance	pitch	timbre	intonation	pitch	volume	timing
MelodicPitchCurve	+		++++		+	+		+
FinePitchCurve	+				++++	++++		
PitchStability	+					++++		
SymmetricStackedHarmonics	++++	+++		+++			+	+++
PitchSymStackedHarmonics	+++	+++	++++	+++	+	+	+	++
SpectroCurve	+++	+++	+	++++				
ArticulationCurve	++++	+++					+++	+++
DynamicsCurve	+++	++++					+++	+
ArticulatedPitchCurve	+++	+++	++++		+	+	+++	++

Table 2.1 List of VEMUS's available graphic curves vs. music dimensions: the more '+' the better the curve is at conveying the music dimension. From (Fober, Letz, and Orlarey 2007)

Fig. 2.14 Pitch, envelope, and harmonics combined curve from (Fober, Letz, and Orlarey 2007)

2.4.3 Learning settings

VEMUS, like IMUTUS, was designed with three learning settings: self-learning, classroom, and distance learning. All three settings benefit from the new functionality added in VEMUS. Furthermore, while the distance learning functionality was proprietary within IMUTUS, VEMUS used the MOODLE (Modular Object-Oriented Dynamic Learning Environment) open source platform (moodle.org 2010) for its enhanced capabilities.

Dima and Borcos (Dima and Borcos 2008) extended VEMUS's distance learning potential further by using the VEMUS platform in conjunction with IBM Lotus Virtual Classroom (In the United States, IBM Lotus Virtual Classroom has since been withdrawn from marketing as of September 11, 2007 and removed from support as of September 11, 2008 (IBM 2010)). The virtual classroom allows students and teachers to connect with each other through a broadband internet connection. When connected, they can make use of the virtual classroom capabilities of real-time audio, real-time video, whiteboard, text chat, and screen sharing. This means that while students are using VEMUS from their home computers, the teacher can remotely view and hear the students' performances in real-time. Furthermore, the teacher can observe the feedback presented by VEMUS through screen sharing. The screen sharing functionality also allows the teacher to share control of students' mice and keyboards, allowing the teacher to add any appropriate annotations.

2.5 DVT

The DVT (Digital Violin Tutor) (Percival, Wang, and Tzanetakis 2007; Yin, Wang, and Hsu 2005) is another computer-assisted musical instrument tutoring system. This system was designed specifically for beginner violin students. It is able to provide visual feedback to students including a transcription of their performances (displayed as a "piano roll" graphical representation), a 2-D fingerboard animation, a 3-D avatar animation, video and audio playback of students' performances, and pre-recorded video and audio of a teacher's

Fig. 2.15 Screen shot of VEMUS with IBM Lotus Virtual Classroom from (Dima and Borcos 2008)

performance (see Fig. 2.16). The system also includes a real-time tuner.

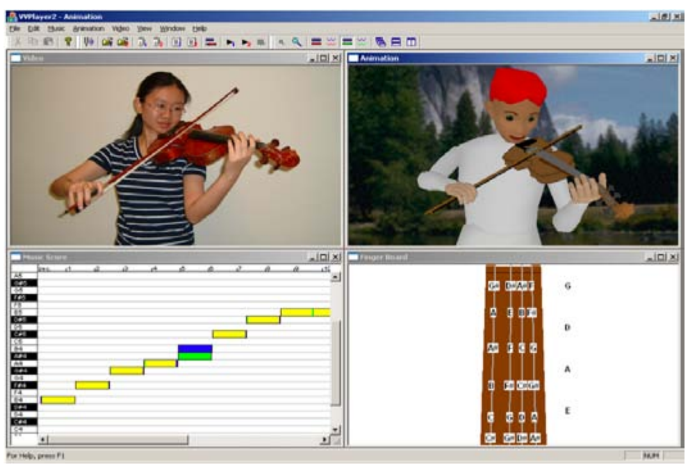

Fig. 2.16 A screen shot of DVT's user interface from (Yin, Wang, and Hsu 2005)

The audio of students' performances is captured using a microphone. It is then transcribed and compared to a transcription of the teacher's performance. This performance evaluation determines the feedback to be displayed. A diagram of the system design is shown in Fig. 2.17.

The 2-D fingerboard and 3-D avatar animation give a visual representation of how

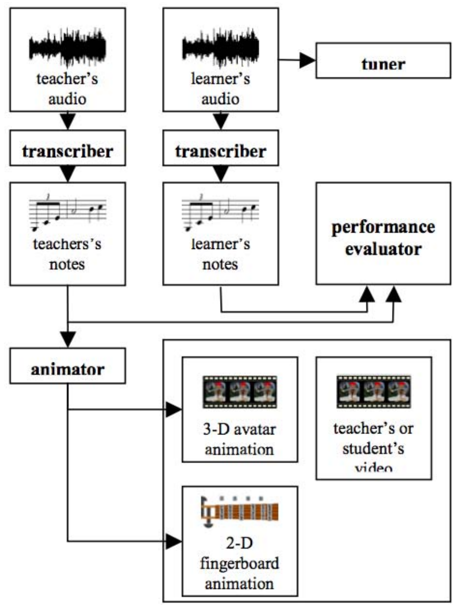

Fig. 2.17 System diagram of DVT from (Yin, Wang, and Hsu 2005)

students should execute specific musical pieces. The 2-D fingerboard highlights the active string and fingering position. The fingering is selected based on the notes read from the teacher's audio transcription. The 3-D avatar animation attempts to give a 3-D version of what a human teacher would demonstrate to students. The 3-D avatar animation is synchronized with the 2-D fingerboard animation. Results from an evaluative study found that the 2-D fingerboard animation was effective in clearly demonstrating the fingering positions, while the 3-D avatar animation was not (Yin, Wang, and Hsu 2005). An indication of how well these graphical images were at communicating intonation preciseness, which is useful in violin teaching, learning and playing, was not specified.

2.6 iDVT

The iDVT (Interactive Digital Violin Tutor) (Lu et al. 2008; Wang and Zhang 2008; Wang, Zhang, and Schleusing 2007) is the extension of the DVT project. Lu, Zhang, Wang and Kheng (Lu et al. 2008) state that it can be difficult to determine the onset detection for instruments like violins which have pitched non-percussive (PNP) sounds. For this reason, the system was designed to amalgamate both audio and visual information to increase the reliability of onset detections. When compared to audio-only results, the results of the multimodal (audio and visual) implementation demonstrated a 10% increase in true positives[5] and an 8% decrease in false positives[6] (Wang, Zhang, and Schleusing 2007).

One microphone is used to record the audio information, while two webcams are used to record the visual information. One of the cameras is positioned facing the front of the performer to capture bowing strokes. The second camera is positioned above the student to capture fingering movements.

The audio and visual data are processed individually for onset detection. The multimodal data fusion process consolidates the results of both. Based on these onsets, the original audio recording is segmented into note segments for pitch estimation. A diagram of the entire process is shown in Fig. 2.18. More details of this process can be found in (Lu et al. 2008).

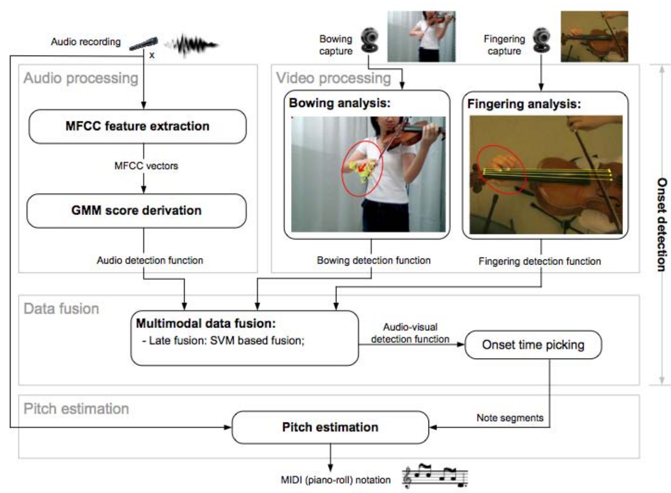

Fig. 2.18 System diagram of iDVT from (Lu et al. 2008)

2.7 i-Maestro

Similar to VEMUS, i-Maestro (Interactive Multimedia Environment for Technology Enhanced Music Education and Creative Collaborative Composition and Performance) (i-Maestro 2010; Ng 2008a,b, 2009; Ng, Weyde, and Nesi 2008; Ng et al. 2008; Ong et al. 2006) incorporates a variety of computer aided technologies to enhance, and not replace,

[5]true positives $= \frac{numberOfCorrectlyDetectedOnsets}{numberOfAnnotatedOnsets}$

[6]false positives $= \frac{numberOfWronglyDetectedOnsets}{numberOfAllDetectedOnsets}$

traditional music tuition. It was designed specifically for bowed string instruments such as the violin, viola, and cello.

Like VEMUS, i-Maestro can be used for self-learning, one-to-one (a student and a human teacher), and/or a classroom environment (multiple students and a human teacher). The i-Maestro system architecture consists of three layers: authoring layer, storage layer, and execution layer (see Fig. 2.19).

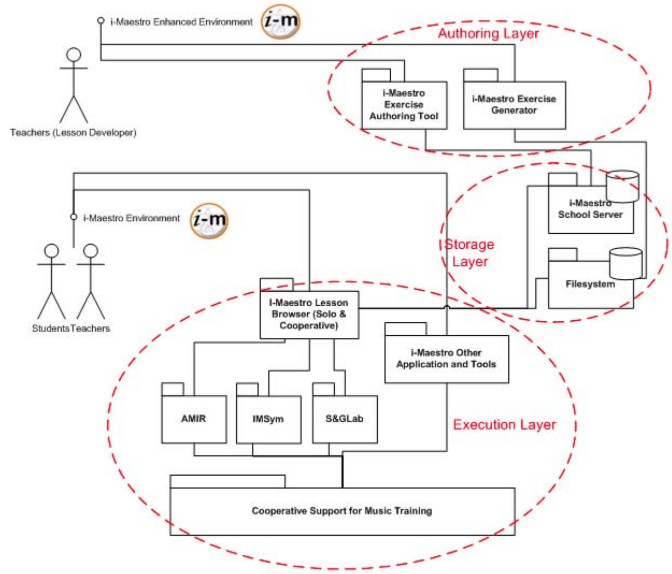

Fig. 2.19 I-Maestro structure from (i-Maestro 2010)

Within the authoring layer, teachers are able to create music exercises and lessons for the students. The Exercise Authoring Tool is a software tool which allows teachers to create learning units consisting of exercises and lessons. The Exercise Generator is a semi-automated tool for generating variations of existing exercises, as well as synthesizing new exercises based on user defined variables. This automation saves the teacher time when creating course material.

The storage layer consists of a School Server. This server is used to manage, store, and retrieve data regarding student progress and history, exercises and lessons, pedagogical material, and user access to the i-Maestro.

The execution layer encapsulates different client tools used by students to enhance their music learning. The tools include the Symbolic Trainer (IMSym), 3-D Augmented Mirror (AMIR), the Sound and Gesture Lab (S&GLab), and Cooperative Learning Tools.

The Symbolic Trainer allows a single user to practice music theory exercises. The exercises may consist of multiple choice questions, editing questions using Music Editing Service for Max, pairing exercises (e.g. students must select which score is associated with an audio track), or ordering exercises (where students must reorganize a musical score that was spliced and shuffled).

The 3-D Augmented Mirror (AMIR) helps teachers and students evaluate performance gestures. Using 3-D motion capture technology, students' bowing and posture gestures are captured and conveyed through visual feedback and sonification. A screen capture of AMIR is shown in Fig. 2.20. Students can greatly improve their playing by being aware of their gestures, postures and technique. This allows students to have a deeper insight into their performance, and can elucidate ways in which they can improve. The reason sonification, rather than visualization, is useful is that students may be visually focusing on other information (e.g. they may be focusing on a score). Therefore, the system uses sonification to communicate details of students' gestures in real-time. For more information on the 3-D Augmented Mirror see (Ng et al. 2007).

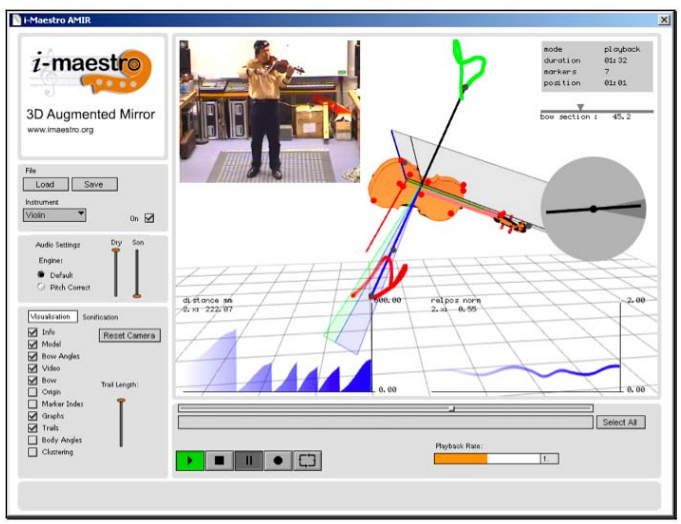

Fig. 2.20 Screen capture of what is seen on AMIR from (Ng 2008b)

Performance gestures are also utilized in i-Maestro's Sound and Gesture Lab (S&GLab). With the use of a wireless gesture sensor attached to the performers' wrists and the technology from the *augmented violin* project (Bevilacqua et al. 2006) bowing gestures can be tracked and utilized for gesture following (see Fig. 2.21). Score following is also offered through real-time audio analysis in reference to a symbolic music representation (SMR) of the performed piece.

These technologies can be used for adaptive accompaniment, where the i-Maestro software will play audio accompaniment which adjusts its tempo according to the tempo provided by students' performances. Furthermore, if the wireless gesture sensor is used by another individual who is conducting, their gestures can also dictate the tempo of the accompaniment.

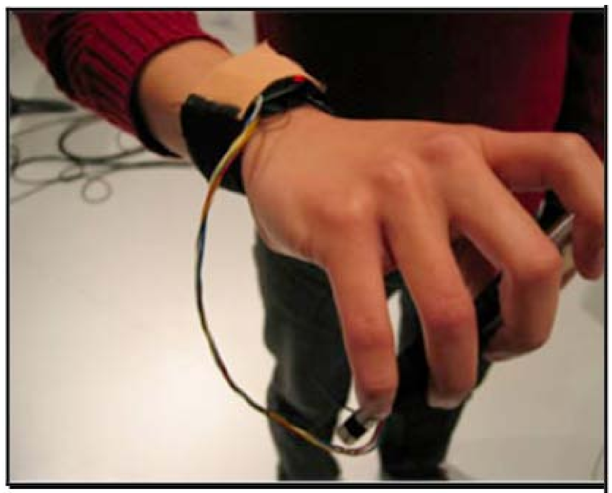

Fig. 2.21 Wireless gesture sensor attached to student's wrist and violin from (Ng 2008b)

i-Maestro's Cooperative Learning features allow students to collaborate in real-time. Cooperative learning activities include ear training exercises and collaborative music authoring. While students perform these activities, any changes made by the students can be seen in real-time by the other students and the teacher. This means as multiple students perform an activity, such as the ear training exercise shown in Fig. 2.22, the teacher will be able to see the students' progress in real-time. By monitoring the students' progress in real-time the teacher will be better able to offer assistance or praise to the students as they work.

2.8 Interactive Software for Guitar Learning

Smith and Johnston (2008) outline a guitar learning software designed to help beginner guitar students. The goals of the interactive software are to improve students' motivation to practice, as well as to improve their musical listening skills by offering feedback and musical

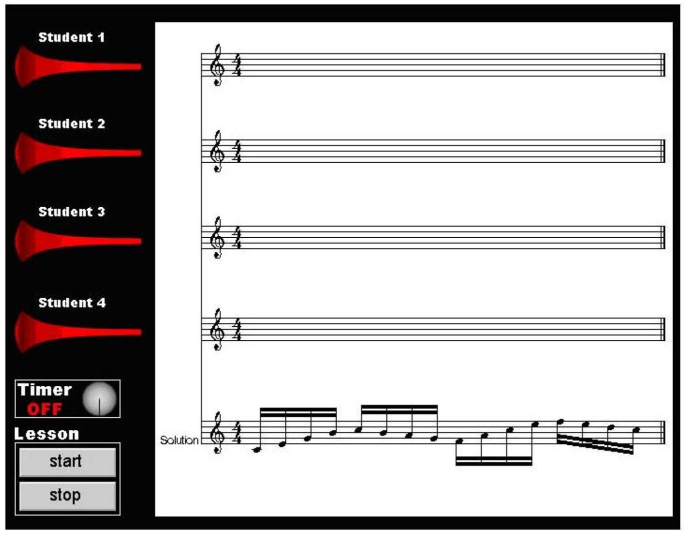

Fig. 2.22 A screen capture of what a teacher may see during cooperative ear training exercise from (i-Maestro 2010)

accompaniment and making students more aware of their performances (e.g. through visual feedback of correctly and incorrectly played notes).

In the implementation of the software, polyphonic pitch estimation was not used. Smith and Johnston state that polyphonic pitch estimation methods are currently not accurate enough to provide reliable feedback to students. Instead, the software implements monophonic pitch detection on the audio signal from the guitar's output jack.

A typical session involves students selecting a song to perform from a list provided by the software. Students can choose to enable or disable MIDI accompaniment as well as a MIDI version of what they are to perform. The software displays the notes students are to play through a side scrolling tablature representation (see Fig. 2.23). As students perform the piece, real-time feedback is shown indicating whether they are playing the correct or incorrect timing and pitch. The feedback is given graphically as a green check mark or a red X, as well as through shading of the tablature area either green or red respectively (see Fig. 2.23).

Feedback is also displayed after the performance is completed. An overall accuracy percentage is shown and the students have the ability to listen to a recording of their performance in order to analyze and review what they did correctly or incorrectly (see Fig. 2.24). A piano roll representation of the performances can be toggled on or off. This visual representation gives students another way of viewing the timing and notes they

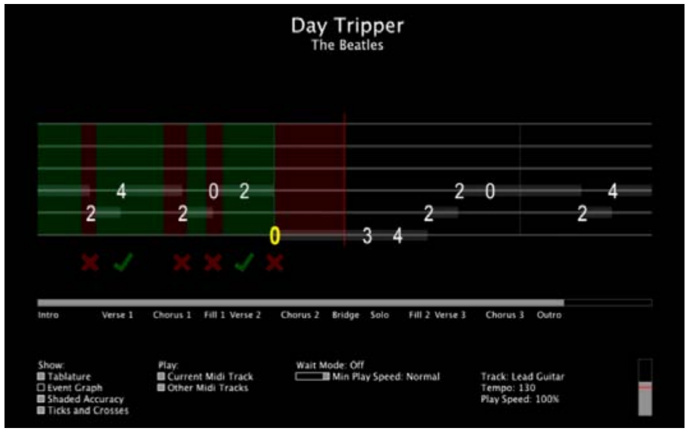

Fig. 2.23 A screen shot of the Interactive Software for Guitar Learning as the student performs and receives real-time feedback. From (Smith and Johnston 2008)

should have played and what they *actually* performed (see the middle area of Fig. 2.24).

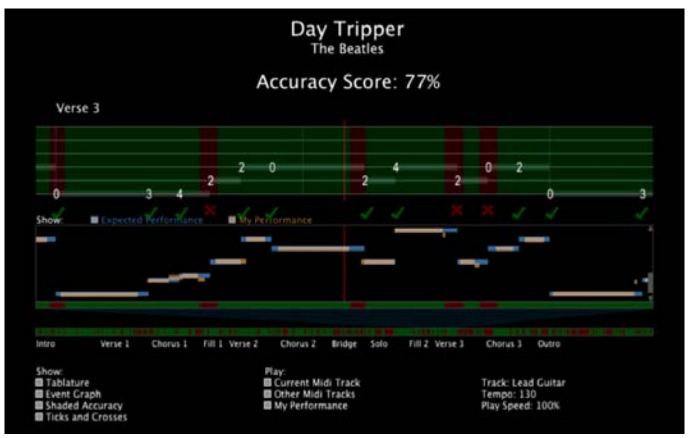

Fig. 2.24 A screen shot of the feedback a student may review after their performance from (Smith and Johnston 2008)

As shown in the bottom region of Fig. 2.23 and Fig. 2.24, the user interface allows students to customize a variety of settings. Students can toggle a display of the tablature view, event graph (piano roll), shaded accuracy, and tick and crosses. Students can also toggle the audio of the current MIDI track (a MIDI version of what they are to perform), other MIDI tracks (MIDI accompaniment), and their performance (when reviewing their performance). Students may also increase or decrease the tempo of the song.

Another noteworthy feature is the Wait Mode (also shown in the bottom of Fig. 2.23).

When playing a song in Wait Mode, the system will wait for students to play the correct note. If they play an incorrect note, Wait Mode gives them the time to focus on correcting their mistake before moving to the next note. Wait Mode is especially useful for beginner students who may not be very familiar with playing the correct notes. These students may require time to look at their hands to figure out the correct fingering before looking back to the screen. Students can customize how much the song will slow down (or even stop) while the system waits for the correct note to be played. A button slider is used to select the minimum slow down speed, which ranges from completely stopping the song to not slowing down at all (in which case Wait Mode is turned off). Wait Mode can be a useful tool for familiarizing students with how to play notes correctly. However, it should not be used exclusively, as it does not allow students to practice rhythm.

2.9 Guitar Hero & Rock Band

The popular game franchises Guitar Hero (RedOctane 2005) and Rock Band (*Rock Band* 2007) brought rhythmic music games to a broad audience. Both game franchises incorporate guitar shaped controllers which players use to play along to popular songs. The guitar controllers consists of five buttons on the neck of the guitar which act as a way for players to simulate fretting the guitar, and a toggle switch on the body of the guitar which simulates picking the guitar (see Fig. 2.25). Players must coordinate pressing the buttons and toggle switch with corresponding graphical circles (representing the notes they are to fret and pick) that descend in time with the musical accompaniment (see Fig. 2.26).

Fig. 2.25 Guitar Hero controller

In addition to using a guitar controller, Rock Band also functions with a custom drum controller, piano keyboard controller and a microphone. Recently, Rock Band 3 has expanded to include two new guitar controllers: the Wireless Fender Mustang PRO-Guitar

Fig. 2.26 Screen shot of original Guitar Hero

Controller (see Fig. 2.27) and the Squier by Fender Stratocaster Guitar Controller (see Fig. 2.28). The Wireless Fender Mustang PRO-Guitar Controller is closer in terms of simulating the experience of playing a real guitar as its guitar neck has 17 frets with 6 buttons per fret and 6 strumming/picking strings by the body of the controller. However, the Squier by Fender Stratocaster Guitar Controller[7] is in fact a real guitar with fret sensors for each fret and string (*Rock Band* 2007). This controller allows players to play a real guitar as the controller for the game. The technology of this controller is discussed further in Sec. 4.3.5.

Fig. 2.27 Rock Band 3 Wireless Fender Mustang PRO-Guitar Controller from (*Rock Band* 2007)

The Guitar Hero and Rock Band games also incorporate the use of avatars, storyline, challenges, feedback, and other game elements which can increase players' engagement in

[7]As of the writing of this thesis (January 24, 2011), the Squier by Fender Stratocaster Guitar and Controller has not yet been released. It is scheduled to be released on March 1, 2011.

Fig. 2.28 Rock Band 3 Squier by Fender Stratocaster Guitar Controller from (*Rock Band* 2007)

practicing and performing songs. The benefits of using these and other game elements within educational software are described in more detail in Ch. 5.

Musicians or music teachers may criticize the musical merit of Guitar Hero and Rock Band by stating that it does not teach players how to play real guitars (with the exception of the new Rock Band 3 controller), read proper musical score notation, play or understand expressive playing, or fulfill other musical elements (see Ch. 3 for a more in depth review of musical elements useful in music tuition). However, although these games may not teach students all the elements of music, they may be able to teach *some* useful concepts related to music such as:

- executing rhythmic patterns and honing a sense of musical timing
- using muscular dexterity to play a musical instrument by pressing the correct buttons at the correct time
- sight reading (even if the notation used in the game is not the same as traditional musical notation)
- memorizing songs
- experiencing the benefits of repeatedly practicing a song (i.e. a player may practice a song to achieve a higher score)
- experiencing, even to a limited degree, what it is like to perform a song
- introducing players to new music, artists, and cultures
- performing music alongside other individuals

I would argue that individuals without prior musical training who play these games and enjoy the experience will gain an increased interest, motivation, and awareness of music. This may occur due to players being exposed to the content of music in an engaging context. They will be practicing and improving the musical concept described in the list above. Even if these games do no exhaust *all* of what is possible from playing a real guitar, more relevant skills are learned than if they had not had the experience. In a paper discussing the educational merits of Rock Band, Nardo (2010) states:

> "...the time has come to acknowledge and encourage some of the educational benefits these games may offer outside of school and in school. Of course they will never replace real instruments, but they are capable of motivating the student to engage in singing and rhythmic activities for hours on end - with friends and family. These skills might easily transfer to the classroom setting if we provide and extend the learning for our students." (Nardo 2010, p.2)

2.10 Summary of Notable Design Features

This chapter has outlined several innovative features of a few tuitional systems. Some of these innovative features have been applied to subsequent tuition software. The list below summarizes the key innovative features described from each section of this chapter:

- Sec. 2.1: Piano Tutor
  - Score following capabilities
  - Expert System to offer feedback
  - Lesson Database
  - System's actions during a lesson: selects lesson from student model (students can also select a lesson), delivers lesson, records students' performances, analyzes performance, provides feedback, provides any remedial lessons, and updates student model
- Sec. 2.2: pianoFORTE
  - Graphical feedback of the musical elements: dynamics, tempo, articulation, and synchronization

- Sec. 2.3: IMUTUS
  - Three learning settings: self-learning, classroom, and distance learning
  - Three phase interaction: explorative phase, practicing phase, and feedback phase.
  - Performance evaluation module which compares students' performances with the ideal performance, detects errors, assigns a grade between 1-5, and gives students feedback regarding three aspects of playing
  - Annotatable score including highlighted regions, graphical images, and text
  - Fingering viewer which shows a 3-D representation of the correct fingering for notes
  - Virtual tuner
- Sec. 2.4: VEMUS
  - Continuation and extension of the IMUTUS project
  - Extended annotation capabilities to include emoticons, hand written capabilities, audio annotations, and real-time audio processing
  - Graphic curves which visually demonstrate the performance of specific musical features (i.e. articulation, nuance, pitch, timbre, intonation, stability of pitch, stability of volume, and timing)
  - Virtual classroom where a human teacher can have a real-time tutoring session remotely with students while they use the VEMUS software
- Sec. 2.5: DVT
  - 2-D fingerboard viewer
  - 3-D avatar animation of finger position
  - Video and audio playback of student and teacher performances
  - Virtual tuner
- Sec. 2.6: iDVT
  - Fusing video and audio onset detection methods to provide more accurate onset detection results, instead of solely using audio data

- Sec. 2.7: i-Maestro
  - Three learning environments: self-learning, one-to-one, and classroom
  - Authoring tools for creating music exercises
  - 3-D Augmented Mirror to capture and display posture and bowing gestures
  - Sound and Gesture Lab which utilizes the *augmented violin* (Bevilacqua et al. 2006) to track bowing gestures and perform tasks such as adaptive accompaniment
  - Cooperative learning where multiple students can interactively work on ear training and music authoring exercises
  - School server to store and retrieve files
- Sec. 2.8: Interactive Software for Guitar Learning
  - Side scrolling tablature view of the music to be performed
  - Real-time feedback indicating if the correct notes were played at the correct time
  - Green and red colour coding representing correct or incorrect note performances
  - An accuracy score is calculated and shown to students when the performance is complete
  - Students can increase or decrease the tempo of the piece to be performed
  - In "wait mode", the system will wait until students play the correct note before requesting the next note (this is not used for rhythmic practicing, it is more suitable for practicing fingering the correct notes)
  - Turn on/off the sound of the accompaniment or the notes which are to be performed by students
  - Students can listen to the recording which is synchronized with the side scrolling tablature indicating correctly and incorrectly played notes
- Sec. 2.9: Guitar Hero & Rock Band
  - Incorporating game elements to increase engagement and motivation (Game elements are described in more detail in Ch. 5)

Many of these features will be incorporated into the high-level requirements of the proposed electric bass tutoring system described in Ch. 6.

For further reading on the use of music education software see (Dervan and McDaniel 2007; Juslin et al. 2006; Miller 2011; Percival, Wang, and Tzanetakis 2007; Puig et al. 2005; Sing & See 2011; Smart Music 2011; So 2008; Star Play It 2011).

Chapter 3

List of Useful Musical Elements for Electric Bass Tuition

One of the goals of the proposed tuition system is to provide feedback to electric bass students regarding their performances. As seen in Ch. 2, music tutoring systems can provide students with feedback for a variety of musical elements (e.g. pitch, dynamics, and rhythmic timing). However, in order to focus more specifically on the electric bass, it is beneficial to determine which musical elements a bass teacher would focus on when working with students.

Five-time Grammy award winning bassist Victor Wooten is a highly regarded electric bass player and music teacher. Jisi (2003) included him as one of the "Millennium Masters" of bass, along with Marcus Miller, Anthony Jackson, and John Patitucci. He has been voted bass player of the year three times by Bass Player Magazine and is the only bassist to receive this award more than once (Wooten 2010a).

Wooten has given many bass clinics and has released a number of instructional videos. In his most recent instructional video *Victor Wooten: Groove Workshop (Instructional DVD)* (Wooten 2008b) and in his fiction book *The Music Lesson: A Spiritual Search for Growth Through Music* (Wooten 2008a) Wooten lists ten musical elements which, from his 44 years of experience playing the bass, he believes bassists should focus on when learning about and playing music. This list consists of:

1. Pitch / Notes

2. Dynamics

3. Tone
4. Rhythm / Tempo
5. Articulation and Duration
6. Space / Rest
7. Phrasing
8. Technique
9. Emotion and Feel
10. Listening

Wooten acknowledges that there may be other elements of music, however these are the ones which he focuses on when teaching music and the electric bass. In his teachings, Wooten emphasizes that bassists can especially benefit from musically exploring the "2 through 10" of the list (i.e. all the elements of the list excluding the "pitch / notes"). He acknowledges that pitch / notes (which, to Wooten, also encompasses harmony, chords, intonation, key signatures, and other aspects of music related to pitch / notes) have value and should not be disregarded. However, as bass players and integral members of the *rhythm* section, he states "we are groove players first, much before we are note players" (Wooten 2008b). He further emphasizes that musical compliments given to bass players focus on their "groove, feel, or pocket" and not their notes (Wooten 2008b). For this reason, while having a knowledge of pitch / notes is important, it is also integral to bassists' success to be familiar and apt with the "2 through 10."

Since the year 2000, Wooten has held the "Victor Wooten's Bass / Nature Camp" (Wooten 2010b). At this camp, Wooten and his co-instructors teach electric bass students about music and nature using this above list of musical elements as a guide. Some of the other instructors at this camp include notable bass players Steve Bailey, Anthony Wellington, Chuck Rainey, Adam Nitti, and Mike Pope, who also use Wooten's list of musical elements during their teaching.

This list of musical elements coincides well with those specified by other music experts (see Table 3.1). In his book *This Is Your Brain On Music*, Levitin (2006) offers a list of fundamental musical elements including loudness, pitch, contour duration (or rhythm), tempo, timbre, spatial location, reverberation, meter, harmony, melody, and space. Zdzinski and Barnes (2002) present a study to develop a String Performance Rating Scale (SPRS) to assist in guiding the assessment of string instrument performances. The results of this study suggest five principle performance factors: interpretation/musical effect (which includes other performance aspect such as style, melodic phrasing and dynamics), articulation/tone, intonation (which includes pitch), rhythm/tempo, and vibrato. In a review on the relationship between vocal expression of emotion and musical expression of emotion, Juslin and Laukka (2003) outline a list of acoustic cues which are used to communicate emotion. This list of cues includes pitch, intonation contour, vibrato, loudness, attack/rapidity of tone onsets, tempo, articulation, pauses/proportion of sound to silence in successive notes, tempo and rhythm variation, and timbre. Finally, the adjudication forms developed by the National Music Adjudication Coalition[1] (NMAC 2011) outline the assessment criteria of tone quality, pitch (which includes intonation), technique (which includes articulation, attacks, accents, and releases), rhythm (which includes tempi and steady pulse), interpretation/musicianship (which includes expressivity and sense of ensemble), dynamics, breath/mallet/bow management (which includes phrasing), and stage deportment.

These elements can be a useful way for electric bass students to focus and improve upon the different aspects of their playing. It may be argued that the quality of students' overall performances may not necessarily be equal to the sum of each individual musical element. In a paper that questions the reliability and musical validity of using segmented assessments, Mills (1991) states:

> "As I leave a concert, I have a clear notion of the quality of the performance which I have just heard. If someone asks me to justify my view, I may start to talk about rhythmic drive, or interpretation, or sense of ensemble, for instance. But I move from the whole performance to its components. I do not move from the components to the whole. In particular, I do not think: the notes were right, the rhythm was right, the phrasing was coherent, and so on - therefore I must have enjoyed this performance." (Mills 1991)

The proposed electric bass tuition system, however, must compartmentalize the different aspects of students' performances, as it does not have the human intuitive ability to

[1]National Music Adjudication Coalition (NMAC) is a partnership with MENC: The National Association for Music Education and NFHS: The National Federation of State High School Associations, and in cooperation with Festival Disney and the Disney Honors

#	(Wooten 2008a,b)	(Levitin 2006)	(Zdzinski and Barnes 2002)	(Juslin and Laukka 2003)	(NMAC 2011)
1	Pitch, Notes	Pitch, Key, Melody, Harmony	Pitch, Intonation, Vibrato	Pitch, Vibrato	Pitch, Intonation
2	Dynamics	Loudness	Dynamics	Loudness	Dynamics
3	Tone	Timbre	Tone	Timbre	Tone Quality
4	Rhythm, Tempo	Rhythm, Meter, Tempo	Rhythm, Tempo	Tempo, Tempo and Rhythm Variation	Rhythm, Tempi, Steady Pulse
5	Articulation, Duration		Articulation	Articulation, Attack	Articulations, Accents, Attacks, Releases
6	Space, Rest	Space		Pauses	
7	Phrasing	Contour	Melodic Phrasing	Intonation contour	Phrasing
8	Technique		Technique		Technique
9	Emotion, Feel		Interpretation, Style	Emotion	Interpretation, Musicianship, Expressivity
10	Listening				Sense of Ensemble

Table 3.1 Correlation between the Victor Wooten's listing of musical elements with those offered by other music experts

provide a qualitative opinion on the music as a whole without first dividing it into smaller components.

For the purpose of this thesis, I will use this list as a framework from which the system will assess and provide feedback of students' performances. A brief description of each musical element is described in this chapter. The methods for obtaining these musical elements from bass performances are described in Ch. 4.

3.1 Elements Related to Pitch / Notes

This element encapsulates aspects of music which are related to the pitch of notes (e.g. pitch, scales, chords, harmony, key signatures, melody, intonation, vibrato). This is the most common musical element to be discussed and focused on when learning the electric bass or other pitched musical instruments.

When an electric bass teacher is focusing on this musical element, the primary issue is ensuring that students are playing the correct notes. Playing a correct note is related to plucking the correct string while pressing down on that string on the correct fret (issues related to the technique of plucking and fretting are described in Sec. 3.8).

It is not very common for electric bass players to play chords, however it is fairly common for electric bass players to play two notes at the same time (e.g. plucking an open low E string while playing the note E on a higher string).

While intonation is an important factor for fretless stringed instruments, intonation is not necessarily focused on as diligently when teaching how to play the fretted electric bass guitar. Factors that affect intonation are the setup of the bass hardware (i.e. bridge, frets, and neck), the tuning of the strings, the amount of finger pressure used to fret a string, the tension applied to the neck with the fretting hand, and instances of 'bending' strings. When teaching the electric bass, it is important to educate bass students of these factors to ensure their intonation is correct.

3.2 Dynamics / Loudness

The musical term dynamics refers to the loudness of the sound or note. The loudness of a sound is related to the psychoacoustic perception of the sound intensity (i.e. amplitude of the sound pressure wave). The level of loudness or softness of a sound is measured in decibels (dB). For humans, perceived loudness is not linear across all frequencies. An "A-weighted" system (measured in dBA) takes this non-linearity into account, and attempts to specify the loudness in terms of what a human would perceive instead of its physical measured intensity.

When an electric bass is played through an amplifier, the loudness of sound can be adjusted by the volume knob on the bass, the volume knob on the bass amplifier, or in the way the bass is played (i.e. playing the string with strong or soft attacks or using other techniques. See Sec. 3.8 for more details regarding bass techniques). When a string is plucked, its sound will decay over time due to the dissipation of the string vibration. One method to cause the string to sustain its loudness and continue to vibrate without re-plucking it is with aid of an EBow (EBow 2011) or acoustic feedback. However the EBow and feedback are not commonly used when playing electric bass.

It is an important skill for bass players to be able to play notes with a consistent and even loudness. While emphasizing notes with dynamic shifts may be musically pertinent at times, a skilled electric bassist will ensure these dynamic modulations do not occur haphazardly. It is also important for bassists to be able to match or change their dynamic level appropriately with the rest of their musical ensemble. Bassists gain this ability by listening to both themselves as well as the rest of their ensemble (see Sec.3.10).

3.3 Tone / Timbre

Timbre is the quality of sound that distinguishes one instrument from another, irrespective of pitch or loudness. For example, it is the different timbres that allow us to identify the sounds produced by a piano, compared to the sounds produced by a trumpet. The term *tone quality* is also used synonymously with timbre. Some adjectives which are used to describe the timbre or tone of a sound include: mellow, rich, covered, open, dull, bright, dark, strident, grating, harsh, shrill, sonorous, somber, colorless and lackluster (Howard and Angus 2001).

The timbre or tone quality is expressed by spectral energy distribution as well as the way in which the spectral energy changes over time (Wessel 1979). The same type of instrument can also produce a variety of tones or timbres. Audi-Graph (Vasser 2010) is a software which allows musicians to visualize the spectral energy of their sound in real-time. Fig. 3.1 and Fig. 3.2 show the spectral content of two different professional saxophone performers playing the same pitch and loudness while producing differing timbres.

Each instrument has its own unique timbre (which allows us to generally identify instru-

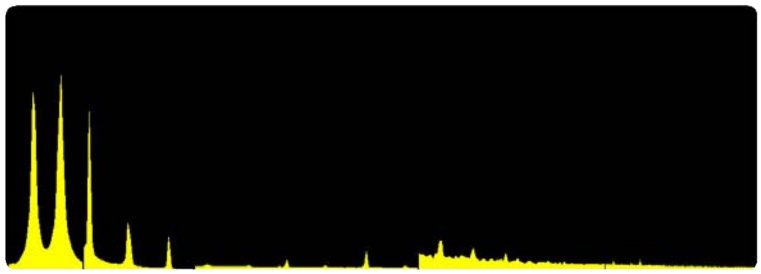

Fig. 3.1 The Audi-Graph display of Dr. Joe Eckert playing an alto saxophone from Vasser (2010)

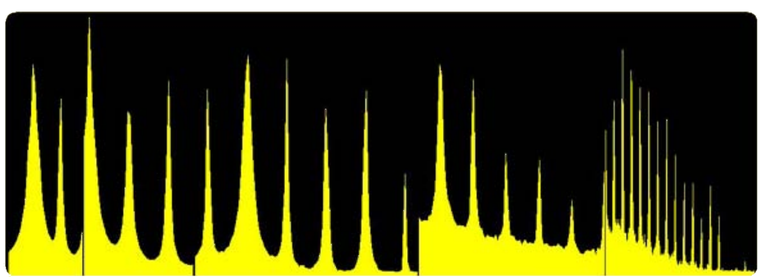

Fig. 3.2 The Audi-Graph display of Trey Gunter playing an alto saxophone from Vasser (2010)

ments) and each instrument can generally produce varying types of timbres. The timbre of the sound produced by an electric bass can be altered based on:

- the brand and model of the electric bass and the materials it is made from
- the type of strings used, as well as how long the strings have been in use (as the oils from players' hands may accumulate on the strings and alter the resonance and tone)
- adjustments to the tone controls (which are variable low-pass filters) and/or toggling between the use of front or rear pickups
- the playing technique used to play the instrument (see Sec. 3.8 for more details on playing technique)
- the type of amplifier and equalizer settings
- use of effect units (e.g. distortion or wah-wah pedals)

When an electric bass teacher works with a student regarding their tone, they would address the factors described in the above list.

3.4 Rhythm, Meter, and Tempo

The term rhythm has varying ways of being defined:

> "Rhythm (in the full sense of the word) covers everything pertaining to the time aspect of music as distinct from the aspect of pitch, i.e. it includes the effects of beats, accents, measures, grouping of notes into beats, grouping of beats into measures, grouping of measures into phrases, etc. When all these factors are judiciously treated by the performer (with due regularity yet with artistic purpose - an effect of forward movement - and not mere machine-like accuracy) we feel and say that the performer possesses a sense of rhythm." (Kennedy and Bourne 2006)

> "Rhythm in music is normally felt to embrace everything to do with both time and motion - with the organization of musical events in time, however flexible in metre and tempo, irregular in accent, or free in durational values." (Latham 2002)

The meter is the overall structure in which the rhythm is organized. The meter is referred to as the time signature. An example of a meter is 3 / 4 (three beats in a bar, and a quarter note is denoted a beat). The tempo is the duration (i.e. how slow or quick) for each beat in the bar (e.g. 120 bpm, represents a tempo of 120 beats per minute).

While accents may be considered a part of rhythm, for the purpose of the list of musical elements, this will be included in *articulation* (see Sec. 3.5).

Bass players are commonly part of the rhythm section of a band, therefore it can be understood that having a good sense of rhythm is vital for a bass player. A bass player, like other musicians, must be able to play both rhythmic patterns specified by the music, but must also know how to interpret these rhythms by adding subtle changes to the rhythm. For example, bassists may focus on playing *on*, *ahead*, or *after/behind* the beat. When bassists play *on* the beat, they attempt to synchronize the attack of their notes as closely as possible with the tempo of the song. When they play *ahead* of the beat, they play slightly earlier. Playing ahead of the beat creates a different sensation, such as the feeling of rushing. Playing *after* or *behind* the beat indicates play slightly later and gives a relaxed or lazy feel.

3.5 Articulation (Attack and Duration)

> "In music an articulation is a sign, direction, or performance technique which indicates or affects the transition or continuity between notes or sounds. Articulations include ties, slurs, phrase marks, staccatos, accents, sforzandos, rinforzandos, and legatos." (Cooper 1985)

The articulation of the note is related to the loudness of the attack and the duration of the note. On a score the articulation of a note is indicated by an accent. A *dynamic accent* is related to the loudness of the attack, while an *agogic accent* is related to the duration of the note. A note which uses a slap or pop technique will have a louder dynamic accent and a short agogic accent. A note that is plucked softly with the pads of the finger and left to vibrate will have an even dynamic accent and a long agogic accent. Other techniques like damping and palm muting can also allow students to alter their articulation. For this reason, the articulation of a note can also provide an insight into what playing techniques bassists are using (see Sec. 3.8 for more details on playing technique).

3.6 Space, Silence, Rests, and Pauses

The terms *space, silence, rests* and *pauses* all refer to the same concept: moments when the performers are not making sound with their instruments. Victor Wooten emphasizes that resting is an effective part of music. Wooten (2008b) divulges “to me, a rest is a note. It’s just one that you don’t audibly hear, but you *feel* it.”

In an excerpt from his book *This Is Your Brain On Music*, Levitin (2006) recounts how musical and visual artists Miles Davis and Picasso make use of space:

> “And as in visual art, music plays on not just what notes are sounded, but which ones are not. Miles Davis famously described his improvisational technique as parallel to the way that Picasso described his use of a canvas. The most critical aspect of the work, both artists said, was not the objects themselves, but the space between objects. In Miles’ case, he described the most important part of his solos as the empty space between notes, the ‘air’ that he placed between one note and the next. Knowing precisely when to hit the next note, and allowing the listener time to anticipate it, is a hallmark of Davis’ genius.” (Levitin 2006, p.17)

An adept professional musician will make use of space and pauses the same way a comedian will pause for just the right moment before delivering a punch line to have the audience in hysterics. In comedy, this is referred to as the comedian’s timing. When joke telling, it is not simply what is said, but how it is delivered. The use of space is an integral part to this, and the same idea applies to music.

Being able to convey feedback regarding space can also help students with their rhythmic timing. There may be a tendency for some bassists to speed up or slow their internal metronome during moments when they are not playing. After resting, this can result in their coming in with notes too soon or too late. Wooten (2008b) states “I have to feel that rest, in order not to rush or slow down.” Therefore, errors in students’ rhythm may in fact be an error in students’ use of space. The two are related, however having a means to communicate space may be an effective way to help students improve.

3.7 Phrasing and Contour

In spoken language, a phrase is a grouping of words which convey some thought or idea. In a paper on the perception of phrase boundaries Knösche et al. (2005) states “phrasing in music denotes the division of the melodic line into structural sub-units, i.e. it is the segmentation of a musical thought for purposes of musical sense." On a musical score, a phrase is sometimes notated with a *phrase mark* which spans from the first note to the last note of the phrase.

An electric bass teacher would need to be able to communicate to students on different ways of phrasing a passage. Therefore the electric bass teacher indicates where one phrase ends and another phrase begins.

A way of distinguishing a phrase boundary is through identifying the ending of a phrase. There is no single unifying method for determining the ending of a phrase. However, Knösche et al. explains "Riemann (1900) considered two cues as most relevant: small breaks immediately after a strong beat and the lengthening of notes, which coincide with a strong beat. In addition, Riemann mentioned several other phrase-establishing cues, including pitch fall and pitch rise in vocal passages or the complete reversal of melodic movement. Harmonic progressions (e.g., cadences and semicadences) have also been described as markers for phrase boundaries." Krumhansl (2000) reiterates this concept: "...longer durations and drops in pitch signal phrase endings–cues that also appear to operate in language. Segmentation can also result from changes in other attributes such as pitch register, timbre, dynamics, and articulation."

Identifying the contour of a musical element may also provide information regarding a musical phrase. A relationship between contour and phrasing was also shown by Huron (1996), who observed how phrases within western folksongs commonly exhibit an arching pitch contour.

Wooten (2008a) suggests that all elements of music can be phrased, not just notes. For example, a recurring consistent series of eighth notes playing the pitch *F#* can exhibit an arching crescendo and decrescendo over the course of 16 eighth notes. This dynamic contour will facilitate the perception of a 16 note phrase. Having the capacity to convey the contour for the musical elements may deepen students' understanding of the musical phrase. Presenting the contour visually could be one way for students to understand the shape of musical phrases.

3.8 Playing Technique

Each musical instrument has its own way of being played. Playing techniques are the methods which are used by musicians to play their instruments. An ideal tuition system would be able to offer the same feedback a human teacher could provide. That would include the system having the ability to provide instructions on executing the techniques as well as gauging students' skill levels in their execution of these techniques. An electric bass teacher would also be able to deduce if students' difficulties in accomplishing satisfactory results with another musical element (e.g. rhythm, tone, etc.) may be due to a lack of proper technique.

In order to understand what may be entailed in determining details of students' technique, Table 3.2 provides a list of common electric bass playing techniques. This is not a fully comprehensive list of techniques, however these are some primary ones which bassists use. For the purpose of this thesis, I will not be going into the details of executing the techniques, as there are countless instructional books and DVDs available regarding the details (Friedland 1996; Pastorius 1985; Pfeiffer 2010; Willis 2002). However, from this list we can extrapolate that the system would need to monitor the placement and movements of the students' arms, hands, and fingers. The three primary methods a system could use to attain this information would be through audio, video, or gesture sensors.

3.9 Emotion, Feel, and Musical Style

When describing this musical element Victor Wooten chooses to use the word *feel* in broad terms and leaves out specifics to allow it to be open to interpretation by musicians (Wooten 2008b). *Feel* can represent the *emotions* expressed through music. It can correlate to the emotional *mood* of the musician or listener, or the general mood which the music conveys (e.g. suspenseful music can be used to foster an eerie mood in a scene of a psychological thriller movie). It can also represent the musical *style* or *genre* (e.g. jazz, rock, classical, etc.) which is being used.

When musicians refer to another musician's *feel*, they are referring to how that musician expresses, approaches, and executes music. A musician may say "That bassist has a great feel. He really plays in the pocket," which would indicate that the bassist may have solid rhythm, articulates well, or be skilled at one or more other musical elements. This would cause the listener to *feel* what that bassist is playing. It should be noted that a musician who has a great feel for one genre of music may not necessarily have a great feel for another genre (e.g. a bassist who has a great feel for country music may or may not have a great feel for punk music).

When helping students improve their bass playing, one exercise Victor Wooten suggests

Hand(s) Used	**Name of Technique**	**Short description of Technique**
Fretting hand	Standard fretting	Pressing down on the string before a fret
	Hammer-ons	Hammering a finger down on a string to initiate the vibration of the string
	Pitch bending	Bending a string with a fretted finger
	Pull-offs	Pulling a fretted finger off a string in a way that causes the string to vibrate
	Vibrato	Wobbling a fretted finger causing the pitch to oscillate slightly up and down
	Sliding / glissando	Sliding a fretted finger up or down the neck while maintain pressure on the string
Plucking hand	Fingerstyle plucking	Plucking a string with the pads of one or more fingers (not including thumb)
	Slapping	Quickly striking a string with the thumb
	Popping / slap plucking	Using the fingers to lift and release a string at a height that will cause the string to hit the fretboard
	Flatpicking	Using a plectrum to pick the strings (A style commonly used by guitarists)
	4-finger classical guitar fingerstyle	Using the thumb and fingers to individually pluck the strings
	Palm muting	Plucking a string while resting the edge of the palm near the little finger against the strings. This dampens the strings
Both or either hand	Two hand tapping	Similar to hammer-ons, however both hands are used
	Harmonics	Plucking an open string (string not being fretted) while gently resting a finger on a node location of one of the harmonic partials
	Artificial / false harmonics	Plucking a fretted string while gently resting a finger on a node location of one of the harmonic partials
	Dampening strings	Gently putting a finger or fingers on an already vibrating string to dampen and stop the vibration
	Ghost / dead notes	Plucking a string that is muted / dampened to such a degree that it only produces a percussive sound

Table 3.2 Listing of Common Electric Bass Playing Techniques

to students is to play their instrument while not focusing on their technique, or any of the other musical elements. Instead, he suggests players focus on the feel, mood, emotions, or ideas they wish to express through the music, and allow other musical elements to be executed by the players naturally and automatically. Wooten uses the analogy that when we speak we do not necessarily focus our conscious minds on every word we say, how loudly we say the words, or the techniques we use to move our lips, tongue, and lungs. Instead, as we talk, we have an idea or feeling we wish to express and the words, dynamics, techniques, and other elements happen naturally. The benefits of using this type of subconscious / unconscious control over the execution of actions rather than using one's conscious control are explored within books by Gladwell (2005) and Lehrer (2009).

3.10 Listening

Listening is a crucial component of performing music. Musicians cannot improve their playing if they cannot hear and identify what they are doing incorrectly. As they perform, musicians can be listening to all nine previously mentioned musical elements. If musicians are playing with an ensemble, they will also be listening to the nine musical elements coming from the other musicians. If the musicians are performing to an audience, they will also be listening and gauging the audience's response. Skilled musicians will be aware of all these sources of musical information and adjust their performances accordingly.

A significant portion of musicians' skill levels of listening can be attributed to how accurately they can detect the nuances of the nine musical elements previously mentioned in this chapter. Musicians' skills at listening may also be attributed to how well they can focus on this task (as they may have the capability / skill, but their focus may be distracted).

One manner of determining musicians' skills at distinguishing the differences in the musical elements could be through employing ear training exercises. For example, exercises could be used to test whether music students can:

- distinguish between different pitches and chords
- be sensitive to changes in the dynamics
- isolate different instruments based on their timbre
- hear when the tempo is speeding up or slowing down

- hear the difference between short and long articulations
- notice when an instrument stops playing
- identify the beginning and end of a musical phrase
- determine which performance techniques were used solely from auditory cues
- distinguish between two similar but different genres of music

The teacher may also attempt to determine students' listening skills by analyzing students' musical performances. Any errors in the performance may be due to either a lack of skill in their execution or in their listening or musical awareness. Some issues students may encounter while performing that may be an indication of a lack of listening may be:

- playing the incorrect notes
- playing too loudly or softly
- playing with a sharp tone when they should be playing with a warm tone
- speeding up and losing the tempo
- failing to play staccato notes
- playing notes during a rest section
- phrasing a series of notes with the incorrect groupings
- not noticing that they are using the incorrect fingerings
- playing a style of music which is incongruent with the accompaniment

It is the electric bass teacher's responsibility to determine whether performance errors are due to an error in execution or a lack in their ability to hear / identify the error. A human teacher may help students improve their listening skills by recording students' performances and having students listen back. The teacher can also point out where students should focus their listening. If students still cannot hear the error, the teacher may simulate an accentuated version of the error to make it more apparent.

Chapter 4

Methods Used to Extract Musical Elements from an Electric Bass Performance

4.1 Introduction to Music Information Retrieval

Music information retrieval (MIR), as its name suggest, is a field of study for retrieving information from music. There are many sub-disciplines within MIR (see McKay (2010) for a thorough list of the different MIR sub-disciplines). One example of an application of MIR is software that can determine the musical genre of a piece of music. MIR is able to accomplish this by extracting low-level features (e.g. amplitude of the signal, spectral centroid, and spectral flux) and high-level features (e.g. chord occurrence, beat histogram) from the music source. These features can then be used within algorithms to distinguish the musical genre (see Tzanetakis (2002) and Aucouturier and Pachet (2003) for methods of determining musical genre).

When analyzing either digital audio or MIDI streams, extrapolating the content from the source can be a difficult task due to the abundance of possible information. Therefore, it is a common practice within MIR to segment the data into sections (e.g. intro, chorus, verse, or even segments of silence). The segments can be determined based on high-level features such as timbre, rhythm and melodic shapes, as well as using Hidden Markov Models and clustering methods (Aucouturier and Sandler 2001; Chen, Lin, and Chen 2004; Foote 2000; Goodwin and Laroche 2003; Levy 2008).

Segmenting the music is also a vital step when performing the MIR task of score following (Raphael 1999). Score following entails tracking the position of a predetermined score in relation to a performance of the score. In order for a tuitional software to compare and evaluate a student's performance of a known piece of music the system must correlate the correct sections of the performance with the relative sections of the score. Orio, Lemouton, and Schwarz (2003), Goodwin and Laroche (2003), Schnell et al. (2006), Levy (2008), and Cont (2010) provide methods to facilitate score following.

4.2 MIDI vs. Audio Sources for Feature Extractions

Two of the main data sources used within MIR include time-stamped events (e.g. MIDI streams) and digital audio (e.g. wav, aif or MP3). Both MIDI and digital audio have benefits and drawbacks for obtaining the musical elements of a performance. MIDI messages are able represente for the majority of musical elements without the need for signal processing. Fig. 4.1 outlines which musical elements can be represented by MIDI messages.

Musical Element	**MIDI Message**
Fundamental frequency	Note pitch and pitch wheel
Dynamics	Velocity
Tone / timbre	NA
Rhythm, tempo, and meter	Note on / note off
Articulation (Duration)	Note on / note off and velocity
Articulation (Attack)	NA
Space, silence, rests, and pauses	Note on / note off
Phrasing and contour	Extrapolated from MIDI messages
Playing technique	NA
Feel and emotion	Extrapolated from MIDI messages
Listening	Extrapolated from MIDI messages

Table 4.1 Musical elements represented by MIDI information

Basic MIDI messages include: note pitch, pitch wheel (bending the pitch upwards or downwards), note on & note off (onsets / offsets), and velocity (a representation of amplitude). With this MIDI messages, the musical elements of articulation, feel and emotion, and listening can be deduced. However, MIDI messages cannot represent the musical elements of tone or technique. (Mckay and Fujinaga 2006) provide a toolkit for extracting multiple features from MIDI streams which is based on McKay's Master's Thesis (McKay 2004).

Digital audio data can be used with digital signal processing to obtain low level features of a musical performance. It is especially useful because it can capture spectral information. The spectral information provides a means to determine low level features such as: power spectrum, magnitude spectrum, strongest partial, spectral variability, spectral centroid, partial-based spectral centroid, partial-based spectral smoothness, compactness, spectral roll-off point, spectral flux, partial-based spectral flux, method of moments, area method of moments, and mel-frequency cepstral coefficients (MFCC) (McKay 2010). These low level features can then be used to establish high-level features and determine qualities of the musical elements. A number of audio toolkits have been developed to extract a variety of audio features (Lartillot and Toiviainen 2007; Mcennis, Mckay, and Fujinaga 2006; Mckay and Fujinaga 2006; Tzanetakis and Cook 1999, 2000).

4.3 Hardware to Obtain Audio, MIDI, and Video Data

Before features can be extracted from a bass performance, the digital audio or MIDI data must first be collected using hardware. The follow section will outline some of the technologies currently available for accomplishing this.

4.3.1 Sonuus B2M - Bass to MIDI Converter

The Sonuus B2M Universal Bass to MIDI Converter (Sonuus 2010) is one device which converts the audio from the output jack of an electric bass to MIDI messages (See Fig. 4.1). However, the Sonuus B2M is limited to converting a monophonic signal and not a polyphonic signal. Although, many bass lines may be monophonic in nature, there are other technologies, discussed in Sec. 4.3.2, which allow the tracking of bass lines which may be polyphonic.

4.3.2 Divided Pickups

One method of performing polyphonic pitch-tracking is through the use of a divided pickup. A divided pickup transmits the audio signal of each string individually. The audio signal from each string can then be processed individually as monophonic signals. The pitch-tracking of these monophonic signals will be more accurate and efficient than if a single polyphonic signal were tracked. The Yamaha B1D (see Fig. 4.2) and the Roland GK-3B (see Fig. 4.3) are two similar divided pickups that are mounted under the strings near the

Fig. 4.1 The Sonuus B2M Universal Bass to MIDI Converter from (Sonuus 2010)

bridge of the electric bass. They both send the divided pickup signals as well as the signal from the bass' built in pickups through a 13-pin cable.

Fig. 4.2 The Yamaha B1D Divided Pickup from (Yamaha Corporation 2010a)

The Graph Tech Ghost Pickup Systems (Graph Tech Guitar Labs 2010) is another type of divided pickups that uses piezo pickups which are integrated into the saddles of the bridge (see Fig. 4.4 & Fig. 4.5). Unlike the Yamaha B1D and Roland GK-3B, which can be used without physically modifying the electric bass, the Ghost Pickup System requires retrofitting the electric bass. The Ghost Saddle Pickups require installing a compatible bridge for the saddles if one is not already installed. Furthermore, Hexpander Preamp Module and 13-pin output jack must be installed, which requires physically modifying the body of the instrument. Fig. 4.6 shows the necessary components of the Ghost Pickup System. On their website, Graph Tech (Graph Tech Guitar Labs 2010) boasts superior tracking capabilities which is presumably due to the string having direct contact with

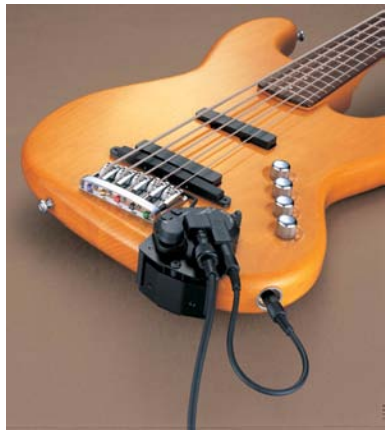

Fig. 4.3 Roland GK-3B from (Roland Corporation 2010a)

piezo saddle.

Fig. 4.4 Graph Tech Ghost Saddle Pickups from (Graph Tech Guitar Labs 2010)

However, for the purpose of allowing many students use of our system, it may be more effective to advise students to mount a Yamaha B1D or Roland GK-3B, as this does not require physically altering their electric bass.

4.3.3 Pitch-to-MIDI Converters

Regardless of the type of divided pickup, a divided signal can then be sent to a pitch-to-MIDI converter via the 13-pin cable. The Yamaha G50 (see Fig. 4.7), Roland VB-99 (see Fig. 4.8), AXON AX 50 USB (see Fig. 4.9), and AXON AX 100 MKII (see Fig. 4.10) are all current technologies which can convert a divided signal to MIDI messages. Further testing would be necessary to determine the comparable reliability and tracking of each system. Each system is able to send out MIDI messages as well as a non-divided audio signal which is transmitted through the 13-pin cable. This non-divided signal could be used in conjunction with the MIDI messages to extract musical features.

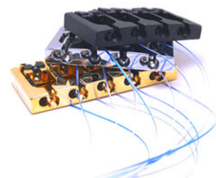

Fig. 4.5 Graph Tech Ghost Saddles mounted into bass bridges from (Graph Tech Guitar Labs 2010)

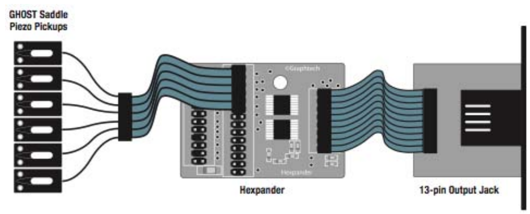

Fig. 4.6 Overview of the Graph Tech Ghost Pickup System from (Graph Tech Guitar Labs 2010)

Fig. 4.7 Yamaha G50 MIDI Converter from (Yamaha Corporation 2010b)

Fig. 4.8 Roland VB-99 V-Bass from (Roland Corporation 2010b)

Fig. 4.9 AXON AX-50 USB from (AXON Technologies 2010)

Fig. 4.10 AXON AX 100 MKII from (AXON Technologies 2010)

4.3.4 Industrial Radio MIDI Bass

A recent technology which has become available is the Industrial Radio MIDI Bass (Industrial Radio 2010). Industrial Radio is an Australian company that manufactures an electric bass specifically for converting its performance to MIDI messages. From their website, Industrial Radio states that the MIDI Bass can attain a latency response of 5 milliseconds when using fixed velocity tracking (not detecting the dynamics) and approximately 8 milliseconds latency when also detecting the velocity (Industrial Radio 2010). These are fairly significant results. These performance results are achieved through the integration of three different sensing technologies. Graph Tech Ghost piezo saddles are used to determine the MIDI note on / off and note velocity. A tension sensing bridge is used to detect the bending for each individual string and translate this to MIDI pitch wheel information. Wired frets are used to determine the MIDI note number/pitch.

The most notable characteristic of the MIDI bass is its wired frets. Each fret is split into isolated segments for each string (see Fig. 4.11). The MIDI bass can determine exactly which string and fret is being played from the contact of a string with a fret segment. This wired fret technology was implemented earlier in the manufacture of the Peavy MIDI Bass and the Peavy Cyberbass (Industrial Radio 2010); both of which have since been discontinued.

Another unique characteristic of the Industrial Radio MIDI Bass is it does not require a separate pitch-to-MIDI converter. Its capacity to convert an electric bass performance to MIDI messages is completely integrated within its system. The entire system consists of the MIDI Bass, MIDI Bass Module, and a specifically designed 8-pin MIDI Bass Cable which connects the two (see Fig. 4.12). The MIDI Bass Module is used as a separate box so there are not multiple wires being plugged in and out of the MIDI Bass. The MIDI Bass Module's input consist of the 8-pin MIDI Bass Cable, MIDI IN, and a sustain pedal (i.e. the same type that would be used to sustain MIDI notes on an electric MIDI keyboard). The MIDI Bass Module's outputs consists of MIDI out, and audio out (i.e. the bass signal from the pickups).

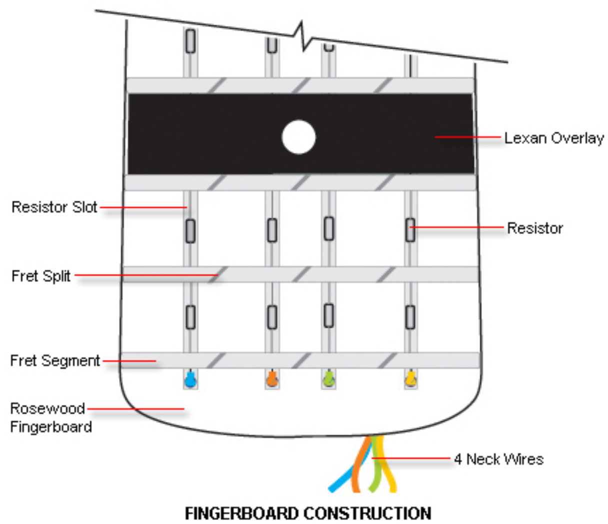

Fig. 4.11 Diagram of the wired frets used in the Industrial Radio MIDI Bass from (Industrial Radio 2010)

Fig. 4.12 Components of the Industrial Radio MIDI Bass system from (Industrial Radio 2010)

The Industrial Radio MIDI Bass appears to be the most effective electric bass to MIDI system, however, at a cost of $ 2495.00 USD it may be too costly an investment for most bass students.

4.3.5 Rock Band 3 Squier by Fender Stratocaster Guitar and Controller

A cheaper facsimile to the Industrial Radio MIDI bass may be developed in the future now that Fender and the makers of Rock Band 3 have developed a real electric guitar with fret sensing capabilities and a MIDI out port. As first mentioned in Sec. 2.9, the Squier by Fender Stratocaster Guitar and Controller (see Fig. 4.13) is a real guitar which has been customized by Fender to be used as a game controller for the video game Rock Band 3 (*Rock Band* 2007). The technology behind this controller is similar to the fret sensing technology discussed in Sec. 4.3.4. There are sensors located in each fret along the fretboard which can detect which fret is pressed for each of the 6 strings on the guitar (see Fig. 4.14).

Fig. 4.13 Rock Band 3 Squier by Fender Stratocaster Guitar Controller from (*Rock Band* 2007)

This controller also comes with a built in MIDI output jack (see Fig. 4.15) which sends out MIDI signals corresponding to what is played on the guitar. For use with game consoles, the MIDI Output signal is connected to the Mad Catz MIDI Pro Adaptor (see Fig. 4.16).

Although this is a guitar controller and not an electric bass controller, the concepts of the hardware and technology could be ported to an electric bass. One of the benefits of this controller is it is priced at $279.99 USD. This is significantly less expensive then the $2495.00 USD of the Industrial Radio MIDI Bass. The price may be significantly less expensive because it does not include the divided piezo pickups or tension sensing bridge which are included in the Industrial Radio MIDI Bass. However, since the Squier by Fender Stratocaster Guitar and Controller has not yet been released[1] the accuracy of the tracking

[1] As of the writing of this thesis (January 24, 2011), the Squier by Fender Stratocaster Guitar and

Fig. 4.14 Guitar Hero controller from (Fender 2011)

Fig. 4.15 MIDI Output jack for the Squier by Fender Stratocaster Guitar Controller from (Fender 2011)

Fig. 4.16 Mad Catz MIDI Pro Adaptor from (Fender 2011)

system cannot yet be determined.

4.3.6 Obtaining Digital Audio Using String Port by Keith McMillen

A benefit of using hardware to convert an electric bass performances to MIDI is it frees up computational resources (e.g. CPU and memory usage). This could be especially useful when the system performs real-time analysis and feedback of a student's performance. It is also beneficial to use any of the described electric bass to MIDI systems because they are self contained systems which calculate the data related to many of the musical elements (i.e. pitch, dynamics, rhythm, articulation, silence). However, the MIDI messages are not able to capture or convey information regarding the tone or playing techniques. Furthermore, the information from the tone and playing techniques are used to understand other details of the bassist's performance. For example, they are necessary for determining the articulated of a note and whether the bassist is playing ghost notes (which are a major asset in a bassist's musical vocabulary).

One possible alternative is to use the divided pickup to route audio data from each string into an audio interface. The audio interface would capture the audio data and the computer system would analyze and compute results of the musical elements. The main limiting factor is whether analyzing all of these musical features would be too computationally taxing on the computer system. Yoo and Fujinaga (1999) determined that the latency of software and hardware methods are fairly comparable. However, conserving computational resources is still important as the proposed electric bass tutoring system will not only be extracting musical elements, it will also be providing feedback about the performance, computing

Controller is scheduled to be on March 1, 2011 graphical images / animations, and computing any other game processes.

A superior hardware alternative, which was implemented by Frisson et al. (2009) for electric guitar, may be a divided pickup which sends its signal to the String Port (see Fig. 4.17) developed by Keith McMillen which costs approximately $795.00 (McMillen 2010). The String Port takes its input from the 13-pin cable that most divided pickups use. The String Port transmits the audio data of each string to the computer through a USB cable. The String Port also converts the performance to MIDI messages. However, this MIDI conversion has been designed for electric guitar and a version for electric bass is currently pending.

Fig. 4.17 String Port created by Keith McMillen from (McMillen 2010)

4.3.7 Features Extraction Using Computer Vision

There has been a series of research on the utilization of video to capture musical features and gestures for stringed instruments. In her Master's thesis, Burns (2006) investigates and prototypes a method of using computer vision systems to recognize the fingering of a guitarist's fretting hand. The system is designed to determine which string, fret, and finger is fretted by the guitarist. Fig. 4.18 demonstrates an overview of the prototype algorithm for accomplishing this task. This type of information could be utilized to determine which notes are played as well as which techniques are used the student. For further details see (Burns 2006; Burns and Mazzarino 2005; Burns and Wanderley 2006a,b).

There has also been research on the use of augmented reality tools to retrieve or display fingering information for both guitar and electric bass (Cakmakci, Bèrard, and Coutaz 2003; Frisson et al. 2009; Kerdvibulvech and Saito 2008). These researches make use of the Augmented Reality Toolkit (ARToolKit) (Kato and Billinghurst 1999). Kerdvibulvech

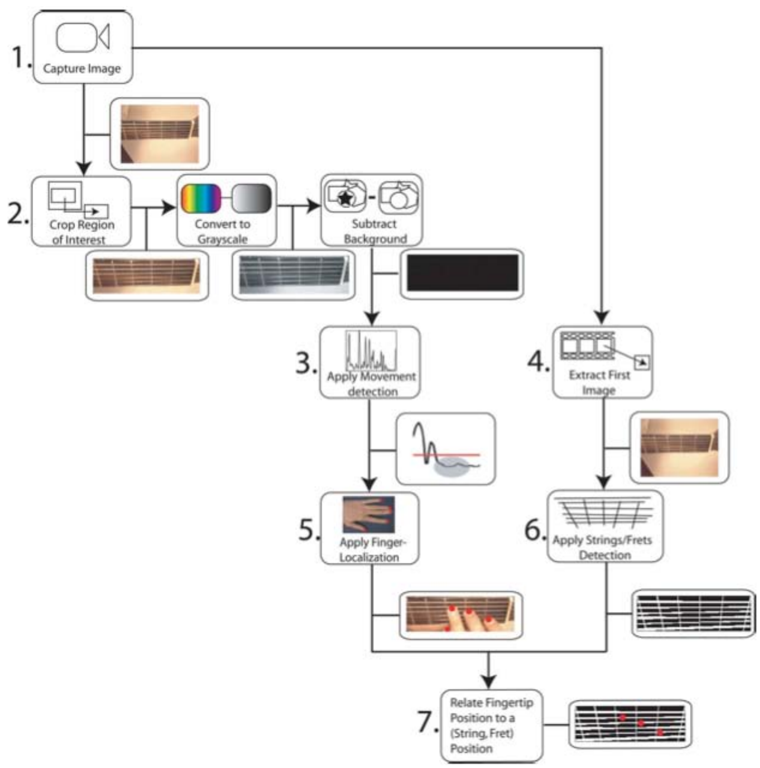

Fig. 4.18 Prototype algorithm for fingering recognition by Burns (2006)

and Saito (2008) take advantage of the ARToolKit to determine finger placement along the fretboard. Users have coloured markers wrapped around their finger tips and two web cams are used to capture videos of the students placing their fingers on the fretboard. Two cameras are used instead of one so the system can extrapolate a 3-D representation, instead of simply a 2-D representation. The system then displays the fingering that it detected. Fig. 4.19 demonstrates the video captured by the two web cameras as well as what the system was able to detect.

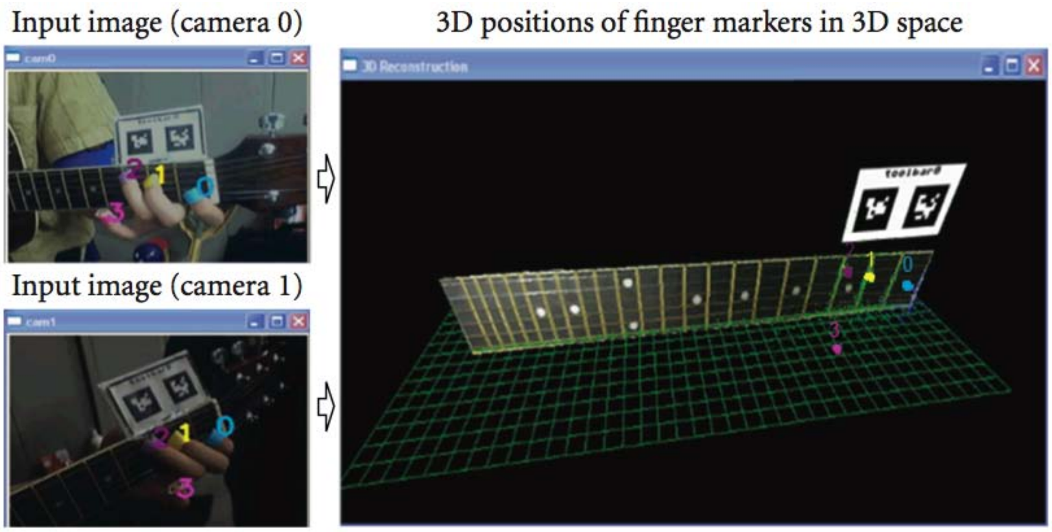

Fig. 4.19 Video from two web cameras and the recognized 3-D fingering position results from (Kerdvibulvech and Saito 2008)

Not only can the ARToolKit be used to assist the system receive information *from* students, it can also assist in the display of information *to* students. Cakmakci, Bèrard, and Coutaz (2003) presented a method for displaying where bass students need to position their finger on the fretboard to play specific notes. The system included a web camera which was aimed at the neck of the bass guitar. A live video of the bass neck was then displayed on a computer screen and the computer software would graphically place a red circle over the appropriate string and fret for the next appropriate note to be played.

4.3.8 Multimodal Approach Amalgamating Computer Vision, Digital Audio, and MIDI Streams

To optimize the accuracy and breadth of the musical element extraction process, a multi-modal fusion of computer vision, digital audio and MIDI streams could be integrated. One difficulty with this method may be that it will also require more computational resources. However, amalgamating the audio, MIDI and computer vision extraction methods will ensure that all the musical elements are accounted for and are accurately extracted. This multimodal approach is similar to the audio and video approach used by the iDVT project (see Sec. 2.6). iDVT was able to combine both audio and video onset detection methods to produce a more accurate onset detection method than by solely using audio data (Lu et al. 2008; Wang and Zhang 2008; Wang, Zhang, and Schleusing 2007).

Frisson et al. presented a multimodal guitar project that consisted of two sub-projects which both aimed at extending the capabilities of the guitar through integrating multimodal technologies. One of the two sub-projects incorporated the work of Kerdvibulvech and Saito (2008), which uses the ARToolKit to determine finger placement on the fretboard. Frisson et al. extracted the fingering position using the methods defined by Kerdvibulvech and Saito and then used this information in conjunction with a real-time polyphonic pitch estimation provided by Cont, Dubnov, and Wessel (2007) to increase the accuracy of detecting which notes were played by the guitarist. Fig. 4.20 displays a screen shot of the fingering tracking results from their implementation. The results are shown as the small red dots near the coloured tips of the fingers.

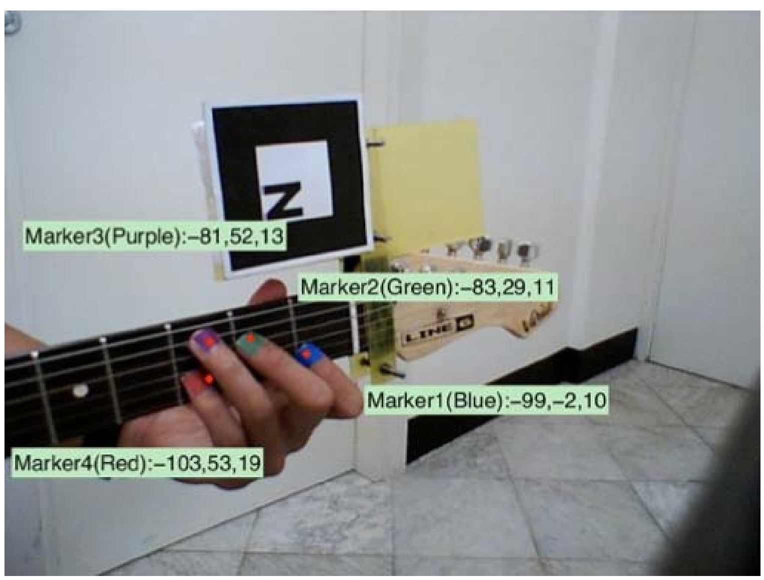

Fig. 4.20 Fingering tracking results from (Frisson et al. 2009)

4.3.9 Suggested Hardware Use for Proposed Tuition System

For this thesis, the multimodal approach of fusing electric bass to MIDI, audio, and video data will be recommended. Retrieving video data can be accomplished with two webcams and the ARToolKit. However, there is currently no hardware which converts bass-to-MIDI as well transmits digital audio for each individual string. The ideal hardware would be have the fret sensor technology used by the Industrial Radio MIDI Bass (see Sec. 4.3.4) and Rock Band 3 Squier by Fender Stratocaster Guitar and Controller (see Sec. 4.3.5) while having a divided pickup which can easily transmit the audio from each string to a computer similar to the String Port (see Sec. 4.3.6).

If the String Port was released for electric bass, it would be the ideal solution as it would be able to send the divided audio signal *and* MIDI data to the computer. However, due to hardware limitations, the optimal hardware configuration would include a divided pickup which transmits to both a MIDI converter, such as the AXON AX 50, as well as an audio interface which can route the divided audio signals of each string to the computer, such as the String Port. In order to facilitate this, the 13-pin connector leaving the divided pickup would need to have a splitter so it may transmit to both the AXON AX 50 as well as the String Port.

4.4 Obtaining Music Elements From a Bass Performance

4.4.1 Fundamental Frequency and Elements Related to Pitch

When obtaining the fundamental frequency of an electric bass performance from an audio signal, it is imperative that students' bass be in tune. Otherwise, students may be playing the correct string on the correct fret, and therefore functionally the correct note, however if that string is not in tune (i.e. too sharp or too flat) the system will perceive this *correct* note as being incorrect.

There have been a number of methods used to obtain the fundamental frequency from an audio signal. Determining the fundamental frequency of a monophonic audio signal has been accomplished fairly successfully. Cuadra and Master (2001) and Gerhard (2003) both provide a review of pitch extraction methods. One widely used method for real-time monophonic pitch estimation is the YIN fundamental frequency estimator (Cheveigné and Kawahara 2002). However, electric bassists are able to play more than one pitch at a time. Therefore, using a monophonic pitch estimation algorithm would not provide accurate or reliable data.

Polyphonic fundamental frequency estimation techniques, on the other hand, are not currently accurate enough to provide valid feedback to students about their performances. There have been attempts at performing multiple fundamental frequency estimation (Chang et al. 2008; Cont 2006; Cont, Dubnov, and Wessel 2007; Klapuri 2006; Sha and Saul 2005). Currently, there is no accurate method of performing polyphonic pitch estimation from an audio analysis of a guitar performance (Smith and Johnston 2008). One successful implementation of polyphonic fundamental frequency estimation appears in Celemony's Direct Note Access (DNA) (Celemony 2010). This technology is able to split a polyphonic signal into individual monophonic signals. The method of achieving this is not currently documented, as it appears to be a trade secret of Celemony's software. However, this technology is not able to split a polyphonic signal into distinct monophonic signals in real-time.

One possible method for resolving this issue of attaining multiple pitch estimations of an electric bass is to combine audio results with other modalities such as computer vision (see Sec. 4.4.9 for more details). Another method is to use a divided pickup (a pickup which outputs the signal for each individual string). With the divided pickup it is possible to isolate the signal of each individual string and then perform monophonic pitch estimation on those signals.

However, the most efficient solution to obtaining pitch from an electric bass is to use technologies which convert bass performance to MIDI messages, as described in Sec. 4.3. If a reliable MIDI representation is used signal processing options would be unnecessary and more computational resources would be available to process the ther musical elements.

4.4.2 Dynamics / Loudness

In MIDI messages, the dynamics are specified by the note's velocity (as previously described in Sec. 2.2.1). While MIDI velocity can be used to indicate the general volume of a note, it does not convey all the intricacies of how a note changes over time. The volume of a note produced by a musical instrument generally follows a contour outlined by its attack, decay, sustain and release (ADSR) (see Fig. 4.21). For this reason, MIDI data is not precise enough to authentically represent the dynamics of a bass performance.

In a paper on the creation of a tuitional software for saxophone players, Robine, Percival, and Lagrange (2007) extract the amplitude from the audio of a saxophone performance by calculating the signal power over small intervals. The signal power is calculated in intervals

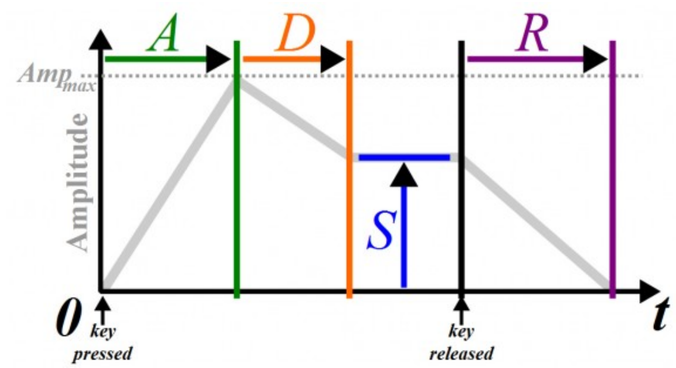

Fig. 4.21 Basic contour of the sound produced by a musical instrument as delineated by its attack (A), decay (D), sustain (S) and release (R). From (*ADSR Envelope*)

because the dynamics of the performance change over time. Fig. 4.22 is the representation of the calculation, where Δ represents the hop size (i.e. the interval size). This method can be used to extract the volume of bass performances.

$$a(i) = \sqrt{\sum_{n=i\Delta}^{i\Delta+I} s(n)^2}$$

Fig. 4.22 Formula for obtaining amplitude from (Robine, Percival, and Lagrange 2007)

A sound level calibration is necessary for a computer system to accurately measure whether students are performing *loudly* or *quietly*. If the system is not calibrated to determine the maximum loudness output of the performance, students may be playing with a lot of force and have their basses' volume knob completely up, however the computer receiving the digital audio may only be receiving seemingly quiet signals, and therefore it will incorrectly deduce the students are playing quietly. Therefore, proper configuration of students' basses, audio hardware, and software setup are important. This may be similar to the sound check bassists would do before live performances or recording sessions. Sound engineers would calibrate the received sound level by requesting the bassists to play as loudly as possible.

4.4.3 Tone / Timbre

Tone or timbre are details concerning the frequency spectral contents of the sound. The spectral content of an audio signal can be acquired first through the use of a Fourier transform. From the spectrum for the signal, further low-level features can be extrapolated such as the spectral centroid, spectral flux, and mel-cepstral coefficients (MFCC). The spectral centroid indicates where the centre of mass of the power spectrum resides. This can be used to generally determine how "bright" or "dark" a sound may be. The spectral flux indicates how much spectral change has occurrs in a signal from one frame to the next. This can be useful for indicating details of how tone changes over time. The MFCC is a feature which groups the magnitude of the spectrum into perceptually organized bins. MFCC is widely used for speech recognition and for MIR because it reflects the human auditory perception system as defined by the mel scale.

Dosenbach, Fohl, and Meisel (2008) provide a method of identifying individual classical guitars by comparing sound samples of the guitars being played. Each sound sample contains a recording of a guitar playing a single note. A support vector machine uses the timbre (functionally using the features of partial tones, Mel-frequency cepstral coefficients, and "nontonal" contributions to the spectrum) and the fundamental frequency of the sound samples to determine which guitar is used. With this in mind, if students have more than one bass they wish to use with the proposed tutoring system, the system could use a technique like this to automatically determine which one is being used. That said, this is likely unnecessary as the system could just as easily have a selection screen whereby students choose which bass they are using.

The tone of a bass leads a listener to perceive its sound as warm, hollow, bright, dark, metallic, dry, etc. In her PhD thesis, Traube (2004) provides an in-depth analysis of the timbre of the classical guitar, which includes a list of verbal descriptors of the timbre of the classical guitar. Although the electric bass will not have the exact same timbre characteristics as a classical guitar, the same verbal descriptors could be used to describe the tone of an electric bass. Other aspects of Traube's work which could be incorporated in the understanding of the timbre of the electric bass is her investigation of how different playing techniques (a.k.a *playing gestures*) affect timbre (Traube 2004; Traube and Depalle 2003; Traube, Depalle, and Wanderley 2003). With the proper investigation, the same concept could be applied to the electric bass.

For additional reading regarding an investigation of how timbre can affect musical perception see the paper published by McAdams (1999).

4.4.4 Rhythm, Tempo, and Meter

The rhythm of an audio signal can be captured by detecting its onsets. Bello et al. (2005) provides a step by step method of onset detection. Using these onsets the inter-onset interval (the interval between successive onsets) can be calculated to establish the rhythm. Within his master's thesis Percival (2008) describes the application of rhythmic exercises which can be used in computer-assisted musical instrument tutoring. With the aid of a light metronome (i.e. a metronome which does not make a click sound, instead it blinks a light for each beat), students would clap a rhythmic pattern which is notated on the computer screen. The system then compares and grades the clapping performance to the expected rhythm.

Data from a video recording of the students performances may also assist in determining the onset of notes. As shown in Sec. 2.6, iDVT demonstrated it is possible to amalgamate audio and video data together to attain more accurate onsets than with audio alone (Lu et al. 2008; Wang and Zhang 2008; Wang, Zhang, and Schleusing 2007). Other ways in which video capture can be utilized in the proposed design of this tutoring system are described in Sec. 4.4.9.

When playing different styles of music, musicians may need to play notes ahead, on, or after the beat. For the system to provide feedback to students regarding their efficiency at executing these subtle beat differences, the system needs to accurately detect the timing of the performers' onsets. For real-time evaluation of note onsets, the system should also take into account any latency that may be due to the technologies used to gather the onset data. For example, if students play a note exactly on the beat, the system may incorrectly detect this note as being behind the beat due to latency. Appropriate testing of this issue would be necessary if the proposed electric bass tutoring system were to be implemented.

There has been a variety of research on beat, meter, and tempo detection (Dixon 2001; Goto 2001; Klapuri, Eronen, and Astola 2004; Scheirer 1998). Although determining the meter of students' bass performances may not be necessary for exercises that already have an established tempo, having the ability to extract the beat and tempo could be useful for exercises that do not have an established tempo. As shown in Sec. 2.7, i-Maestro was able to allow students to increase and decrease the tempo of a performed piece to allow more expressive playing. i-Maestro incorporated a wireless gesture sensor to facilitate this, however it may be possible to use beat detection algorithms which use audio signals as their data source.

4.4.5 Articulation (Attack and Duration)

The articulation students use during a performance can be determined through an analysis of dynamics of the attack and duration of a note. For example, two similar rhythms may have the same inter-onsets, however the duration each note is held during those inter-onsets will determine the articulation that was used for the note. The rate of attack will also have an effect on the articulation of a note. Therefore, to determine the articulation, the system would need to calculate the duration between the onset and offset of a note, as well as detect the rate of attack for the note.

4.4.6 Space, Silence, Rests, and Pauses

Determining moments when students are not playing is not entirely a simple process. Li et al. (2001) outlines a method for pause detection when segmenting a general audio signal. The audio signal is first marked frame-by-frame as either silence or a signal. Each frame is designated as silence or a signal depending on whether the features pertaining to the energy level of the signal are within set threshold boundaries. The relationship between the threshold values and energy level related features are shown in Fig. 4.23.

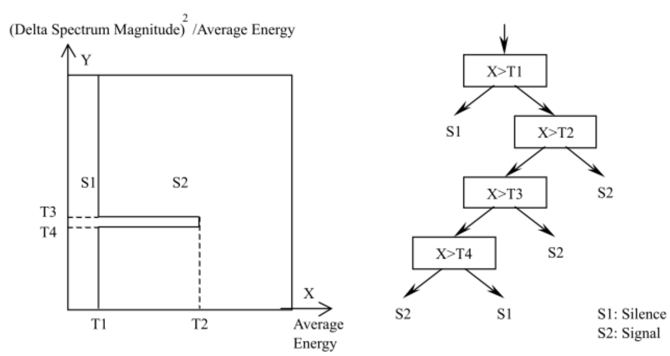

Fig. 4.23 Two-dimensional partitioned space and its corresponding decision tree. Where T1, T2, T3, and T4 are threshold values. From (Li et al. 2001)

The series of frames are then analyzed to check for erroneous signal or silence frames. A silence segment, which represents a continuous series of silence frames, is relabelled as a signal segment if the length is less than a specified fill-in threshold. A signal segment, which represents a continuous series of signal frames, is relabelled as a silence segment if the strength value of the signal is smaller than a specified strength threshold value. The strength value takes into account the fact that powerful bursts of signal energy should not be removed. Fig. 4.24 shows the steps of the process.

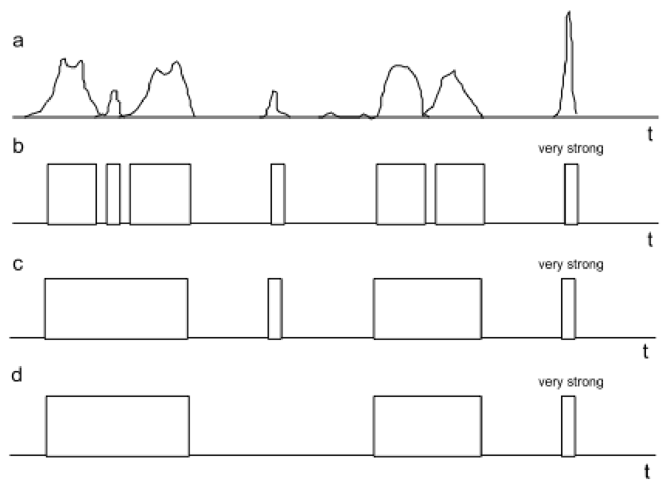

Fig. 4.24 Pause detection steps from Li et al. (2001): (1) short time energy of input signal; (b) initial result for candidate signal segments; (c) result after fill-in step; (d) result after throwaway step.

4.4.7 Phrasing and Contour

As described in Sec. 3.7, some cues which indicate a phrase boundary are:

- small breaks immediately after a strong beat
- lengthening of notes, which coincide with a strong beat
- pitch fall or pitch rise
- reversal of melodic movement
- harmonic progression (e.g. cadences and semicadences)
- the contour of the pitch, dynamics, timbre, rhythm, and articulation

Phrase boundaries could be established by comparing students' performances with the above conditions. Pitch, timbre, dynamics, rhythm, and articulation would be used as the data to match against these conditions.

Automatically segmenting music into distinct phrases is a complex process. There have been attempts at automatically extracting phrases through estimating phrase boundaries and using pattern detection algorithms (Chen, Lin, and Chen 2004; Lu and Zhang 2003; Yanase and Takasu 1999). However, since our system will be working with beginner students, they may perform minor errors which can throw off the accuracy of the phrasing extraction methods. For this reason, it may be more useful to not explicitly estimate the phrasing but to instead display the contour of the pitch, timbre, dynamics, rhythm and articulation. Graphical representations of these contours may assist the players to conceptualize the phrasing of a piece. Examples of displaying this information are the graphic curves that were incorporated in VEMUS (see Sec. 2.4).

4.4.8 Attaining Playing Techniques From Digital Audio

As described briefly in Sec. 3.3, the audio of a guitar performance can provide information regarding the plucking location of a string (Traube 2004; Traube and Depalle 2003; Traube, Depalle, and Wanderley 2003). The work of Traube, while focusing on the guitar, would be applicable in determining the plucking location, as well as the types of techniques, used to pluck the string. By determining the timbral characteristics, especially the brightness, of a note, it may be possible to not only determine the plucking location but also the type of technique used to activate the string vibration (e.g. by slapping, plucking, or tapping the string). Research in the sound synthesis of guitar or bass (see Evangelista and Eckerholm 2010; Rank and Kubin 1997) may also help further identify key acoustical features which may assist in the identification of playing technique.

Abeßer, Lukashevich, and Schuller (2010) presented a successful approach to identifying bass techniques from digital audio. The approach extracts audio features related to 5 plucking techniques and 5 expression styles. The plucking techniques include: finger-style, picked, muted, slap-thumb, and slap-pluck. The expression styles included: normal, vibrato, bending, harmonics, dead-notes. When compared to the techniques listed in the previous chapter (see Table 3.2), the only techniques which were not included by Abeßer, Lukashevich, and Schuller were: hammer-ons, pull-offs, sliding, and dampening strings. However, it may be possible to use this approach to identify these techniques as well.

In their approach, specific salient features were selected to assist in the identification of each technique. For example, to identify the slapping and popping technique, which is characterized by its percussive bright sound, the *spectral crest factor* and *spectral centroid* were chosen because they are effective at measuring the percussiveness and brightness of sound respectively.

A data set of 4300 notes were recorded by 3 basses which corresponded to a mix of the 5 plucking techniques and 5 expression styles. 90% of the recordings were used as a training set and the remaining 10% were used as a test set. From these recordings, the features were extracted, mapped to a feature space, and then classified. Pre-processing steps were implement to reduce the dimensionality of the feature space before classification. Multiple classification methods were also used to find to identify which ones produced optimal results. The best results reached an accuracy of 93.25% and 95.61%. For more details see (Abeßer, Lukashevich, and Schuller 2010).

4.4.9 Attaining Playing Techniques Using Computer Vision

Using computer vision techniques, it may also be possible to determine gestures of the students' plucking hands. This could be accomplished using the Augmented Reality ToolKit with two web cams to achieve a 3-D representation of plucking gestures. The system may then be able to determine what type of plucking technique is being used (e.g. fingerstyle or slapping), as well as where along the string it is being executed.

When the system is able to determine students' playing gestures, it will be able to provide more accurate feedback regarding whether proper or improper techniques are being performed. If the system is able to use some form of computer vision to detect students' posture and gestures (i.e. in terms of their backs, shoulders, arms, hands, and fingers) the system could possibly relay vital information on improving technique to improve playing and to prevent future pain or injury. Computer vision methods may be able to facilitate this. Another solution would be to implement the 3-D motion capture technology that was used by i-Maestro's 3-D Augmented Mirror to capture bowing and posture gestures (see Sec. 2.7).

Another method of helping students improve their techniques is through self assessment. By watching a video recording of their performances, students can observe their techniques and postures from an outside perspective. This different perspective on their performances may lead to valuable insights into aspects of which they may not otherwise be aware.

The system can also help facilitate self assessment by giving students verbal instructions regarding gestures and postures. Exercise or yoga instructors may give verbal instructions to their students regarding certain muscle groups or body movements they should be focusing on. These verbal cues draw students' attention to that particular part of their bodies so they can self correct. This is incorporated in the exercise software *Yourself! Fitness*

(*Yourself! Fitness* 2004). Within the software, a virtual personal trainer named Maya offers verbal suggestions regarding proper form throughout the exercises (e.g. Maya may say "Bend those knees," or "Don't forget to breathe."). This is especially effective because the software is unable to see the actual body gestures of the students.

The incorporation of verbal suggestions or feedback could be used in the design of our system. However, if they are used the verbal cues must not overwhelm or distract students from the actual tasks they are to complete.

4.4.10 Emotion, Feel, Mood, and Musical Style / Genre

For a system to be able to distinguish if students are playing a bass line well, it would be pertinent for the system to have information about what style or genre of music students are attempting to play. Levitin (2006) describes part of the way we identify or categorize musical genres:

> "Each musical genre has its own set of rules and its own form. The more we listen, the more those rules become instantiated in memory. Unfamiliarity with the structure can lead to frustration or a simple lack of appreciation. Knowing a genre or style is to effectively have a category built around it, and to be able to categorize new songs as being either members or nonmembers of that category - or in some cases, as 'partial' or 'fuzzy' members of the category, members subject to certain exceptions." (Levitin 2006, p. 233)

If a system is to provide feedback to assist bassists to improve their skills at playing particular styles or genres of music, the system would need to know what *rules* are related to that style or genre. There may be ways of phrasing or articulating a bass line for one style of music which may be different from another style. For example, a reggae bass line may be laid back and utilize space, while a death metal bass line may use distortion, loud dynamics and fast aggressive rhythms. However, there are always exceptions to the rules.

Aucouturier and Pachet (2003) provides an overview of some methods used to identify the genre of music. However, some of the methods described by Aucouturier and Pachet include the use of general instrumentation to help in the genre recognition, which would not be pertinent to this application (as only an electric bass would be used). Therefore, the genre recognition methods by the system (or more specifically, the ability of the system to determine whether students are playing a musical genre or style correctly) would instead focus on the musical elements previously outlined in this chapter (i.e pitch, dynamics, timbre, rhythm / tempo / meter, articulation / duration, space / rests, phrasing / contour, and playing techniques).

The use of these musical elements is also helpful in determining the emotional expression of a musical performances. The *Feedback-learning of Musical Expressivity (FEEL-ME)* project (Juslin 2004; Juslin et al. 2006) attempts to offer music students automated feedback regarding their ability to convey emotional expressions within their performances. FEEL-ME is able to extract and analyze musical elements (referred to as *acoustic cues* in the papers) such as tempo, loudness, timbre, articulation, and attack to determine the emotional intent of the performers. When using FEEL-ME, students record the performance of a melody with pre-defined emotional intent (e.g. happy, sad, angry, or tender). The system extracts the acoustic cues from the students' performances. From these extracted cues, the FEEL-ME software program displays feedback of the performances through details of how well those cues were executed in communicating the desired emotional expression. Finally, students would record another performance with that feedback in mind. The feedback results of the second attempt would be displayed, and could be compared to the first attempt. The findings from the FEEL-ME project demonstrated that a software program can be effective in improving students' expressive skills (Juslin et al. 2006).

However, when compared to receiving feedback from a human teacher, students were less receptive to feedback from computer software regarding emotional expression. This may be rooted in the idea that computers cannot truly understand the complexities of emotions because they cannot experience them, and therefore their feedback is not as justified. The resistance students had to the computer feedback regarding emotions in comparison to human feedback is studied further by Karlsson, Lindström, and Juslin (2009).

Future experiments would be necessary to apply the methods of Aucouturier and Pachet (2003) and Juslin et al. (2006), specifically to the electric bass.

4.4.11 Listening

As described in Sec. 3.10, musicians' listening skills are defined by the ability to focus their awareness on different musical elements, as well as the ability to distinguish details and differences within a musical element (e.g. differentiating between minor or major chords).

There is no specific way to directly identify whether performance errors are due to a lack in students' listening abilities or a lack in their execution of the musical materials. One tactic could be to include student models which track students' abilities over time. Using this student model, when the system detects students have made an error in their performances, the system can correlate that error to whether they should have the ability to be able to *hear* what they are doing incorrectly and then focus on improving the execution of that particular error. The more data it acquires regarding students' skills the more it can establish an approximation of students' abilities to distinguish and identify the differences for each musical element.

Another method to assist students may be to offer them the choice of hearing how the erroneous musical passage should sound. This will improve students' abilities to internalize how the passage should be executed. For example, when the system analyzes students' performances and discovers a section where students played an incorrect rhythm, the system could offer students the option of hearing what they played and how it should have sounded.

The system could also illustrate students' errors through graphical means. Students can be shown what they played and what should have been played. This concept of improving students' listening through giving visual representations was used by piano FORTE (see Sec. 2.2), VEMUS (see Sec. 2.4), and the Interactive Software for Guitar Learning (see Sec. 2.8).

4.5 Summary

This chapter has explored different research, methods, and technologies used to extract and analyze the musical elements of an electric bass performance. They allow the proposed electric bass tutoring system to detect and provide feedback regarding these musical elements.

For this thesis the suggested implementation for extracting the musical elements is a multimodal fusion of digital audio, MIDI streams, and computer vision. As specified in Sec. 4.3.9, there is currently no one piece of hardware which is able to capture both MIDI and audio from a bass performance. With the currently available technologies the suggested hardware to facilitate this includes:

- An electric bass
- A computer system (with keyboard, mouse, and monitor)
- Two web cams and the ARToolKit marker
- A divided pickup (e.g. Yamaha B1D or Roland GK-3B) and a dual split 13-pin cable
- The String Port by Keith McMillen (McMillen 2010) with USB cable

- The AXON AX 50 USB (AXON Technologies 2010)

In order for the system to obtain accurate and useable data, it would need to test and ensure the bass' strings were in tune (pitch accuracy), the volume levels were checked (dynamics accuracy), and the latency of each hardware device was calculated (onset and offsets accuracy).

The *pitch / fundamental frequency* could be attained from audio sources using a divided pickup and the YIN fundamental frequency estimator (Cheveigné and Kawahara 2002), or through using fingering data attained from computer vision (Burns 2006; Kato and Billinghurst 1999; Kerdvibulvech and Saito 2008). However, the most efficient method would be to use a reliable method of bass-to-MIDI converter to acquire the pitch data.

The *dynamics / loudness* could be attained from the MIDI velocity data, as well as the amplitude of the audio signals.

The *tone / timbre* could be attained and conveyed to students' using concepts applied to the study of the timbre of the classical guitar by Traube (2004).

The *rhythm, tempo,* and *meter* can be attained from the MIDI note on and note off data, as well as the onset and offset detection from the audio signal. Beat and tempo extraction could be attained through the methods described by (Dixon 2001; Goto 2001; Klapuri, Eronen, and Astola 2004; Scheirer 1998). Evaluating rhythmic accuracy could be done using the concepts from the rhythmic exercise software created by Percival (2008).

The *articulation* can be determined by using data from the *dynamics / loudness* as well as the onset and offset detection from both MIDI messages and/or audio signal.

The *space, silence, rests,* and *and pauses* can be determined from the duration between the offsets and onsets from both MIDI messages and/or audio signal.

Automatically determining *phrasing* of a performance may not be accurate enough to provide students with reliable feedback. However, the changing *contour* of musical element such as pitch, dynamics, timbre, rhythm, and articulation can be accomplished by the system and then conveyed to students through graphic feedback.

The *playing technique* could be acquired by extracting and mapping salient audio features as described by Abeßer, Lukashevich, and Schuller (2010). This data could also be fused with the results obtained by the computer vision techniques incorporated by Burns (2006) and by those who used the Augmented Reality Toolkit (Cakmakci, Bèrard, and Coutaz 2003; Frisson et al. 2009; Kerdvibulvech and Saito 2008). Students can also review a video recording of their performances to encourage self assessment.

The *emotion, feel, mood,* and *musical style / genre* may be attained by applying the concepts provided by Aucouturier and Pachet (2003) and Juslin et al. (2006).

The system may not determine whether a student is *listening*. However, the system can help improve students' listening skills by offering ear training exercises and auditory playback. Providing students with graphical representing what they hear may also improve their listening skills. This method of offering a graphical representation of their performance is one of the benefits of using a computer as the medium for tuitional communication. The use of graphical representations of sound was also incorporated also by pianoFORTE, VEMUS, and the Interactive Software for Guitar Learning.

Chapter 5

Game Elements Leveraged in Digital Game Based Learning

The use of technology to learn and practice does not necessarily result in a more enjoyable learning experience or increase a student's desire to utilize the technology (Kiili 2005). While it is possible for educational software to teach students how to play a musical instrument, if the game does not engage them, they will not have the drive to continue using the system.

For engaged learners however, video games can be a very useful medium (Dondlinger 2007; Garris, Ahlers, and Driskell 2002; Gee 2003; Kiili 2005; Prensky 2001; Reeves and Read 2009; Shaffer et al. 2005; Squire 2006; Wilson et al. 2009). An interest in playing the game results in more focused attention and a desire to play it again.

Ch. 2 provided a survey of current technologies and methodologies used by different tutoring systems to acquire, analyze, and offer feedback to students about their musical performance. However, all of them, with the exception of Guitar Hero, neglect to structure an interactive experience which utilizes the game elements that have been shown to increase students' engagement.

In the same way that Ch. 3 outlined the key musical elements, this chapter will outline the key digital game elements that are important in creating an effective electric bass tutoring system.

Throughout this chapter, I will refer to prominent authors on the subject of video game design and its effectiveness as applied to learning. Below lists a brief summary of the topics discussed by each author in the referenced articles or books:

- Garris, Ahlers, and Driskell (2002), Dondlinger (2007), and Wilson et al. (2009) provide thorough summaries of the correlation between video game elements and their effective application to educational goals

- Kiili (2005) presents a gaming model that is based on experiential learning theory, flow theory and game design. Kiili's model especially focuses on the factors attributed to Csikszentmihalyi's (1990) theory of *flow*. The concept of flow is discussed further in Sec. 5.4

- Prensky (2001) presents a framework for successful digital game-based learning

- Reeves and Read (2009) explore and describe how game features from digital games can be applied to improve business and work collaboration

- Gee (2003) discusses how digital games, which may not have been designed specifically for educational purposes, nevertheless produce positive educational results. Within this book he describes his notable 36 learning principles in video games

- Kim (2009, 2010), who has helped design successful games such as *The Sims*, *Ultima Online*, and *Rock Band*, outlines how game mechanics and metagaming can be leveraged to increase engagement, especially in social communities

- Two leading books on game design by Rollings and Adams (2003) and Salen and Zimmerman (2003) offer thorough explanations about how to successfully design digital games

These all authors all discuss what they believe to be vital game elements for education, or in some cases simply "good game design." While there is no consensus on a definitive or irrefutable list of beneficial and necessary game elements, I have synthesized a list of the key game elements that have been deemed effective in engaging learners. These elements and their proponents are shown in Fig. 5.1.

Authors who advocate the usefulness of a particular game element are indicated using a check mark. If authors are not checked off under a element, they may or may not advocate its usefulness. This is because there was either no significant mention of that game element in the cited publication or not enough to assume its importance to the author. It should also be noted that this is not a comprehensive list of *all* the elements that can or should be implemented in an educational game. Such a list would go beyond the necessary scope of this thesis. Rather, the purpose of this list is to point out the game elements that appear

Game Elements	**(Dondlinger 2007)**	**(Garris, Ahlers, and Driskell 2002)**	**(Gee 2003)**	**(Kiili 2005)**	**(Kim 2009, 2010)**	**(Prensky 2001)**	**(Reeves and Read 2009)**	**(Rollings and Adams 2003)**	**(Salen and Zimmerman 2003)**	**(Wilson et al. 2009)**
Avatar			✓		✓		✓	✓		✓
Game Worlds & Fantasy	✓	✓	✓	✓		✓	✓	✓	✓	✓
Story & Narrative	✓		✓	✓		✓	✓	✓	✓	✓
Flow		✓	✓	✓	✓	✓	✓		✓	✓
Control	✓	✓		✓	✓		✓	✓	✓	✓
Goals & Rules	✓	✓	✓	✓		✓	✓		✓	✓
Challenge & Game Balance	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Feedback & Rewards		✓		✓	✓	✓	✓			✓
Social Interaction	✓		✓		✓	✓	✓	✓	✓	✓

Table 5.1 List of engaging game elements

to be most useful for the proposed electric bass tutoring system. The integration of these game elements into the proposed design are described throughout Ch. 6.

Some of the game elements in this list emphasize on capitalizing on the use of online digital games. Because of the internet, multiplayer online games and social media web sites have become increasingly popular. Prensky (2001) comments on the video game industry's shift towards multiplayer games:

> "Despite the industry's initial (prenetworking) focus on single-player games or games played against the machine (an era in which we are still involved), just about *all* of today's computer games have become multiplayer in one form or another" (p.123).

Finally, both Rollings and Adams (2003) and Salen and Zimmerman (2003) refer to Bartle's (1996) essay "Hearts, Clubs, Diamonds, Spades: Players Who Suit MUDs" to explain the different ways players approach online games. In his essay, Bartle outlines four different playing approaches to *multi-user dungeons* (MUDs). Although the essay was geared towards MUDs, the principles can be applied to other multiplayer online games. Each playing approach is outlined below:

- Hearts (Socializers) enjoy socializing with others (see Sec. 5.5)
- Clubs (Killers) enjoy imposing upon others (see Sec. 5.5)
- Diamonds (Achievers) enjoy attaining achievements (see Sec. 5.4.4)
- Spades (Explorers) enjoy exploring the game world (see Sec. 5.2)

Players use a combination of these approaches, however, one is generally more dominant (Salen and Zimmerman 2003). When designing an online multiplayer game it is important to appeal to each of these player types equally. Salen and Zimmerman (2003) explain that recounts "...a healthy MUD community required a certain proportion of each of these types of players and adjusting the game design to attract or discourage any given type of player would tend to influence the numbers of others as well" (p.522).

5.1 Avatar

According to Rollings and Adams (2003), there are two main perspectives that video game players have when interacting with the game world (see Sec. 5.2 for more details on game world). The first is as an *avatar* who is a controllable character that exists in a single place in the game world and is limited to performing actions in that space. The second is an *omnipresent* player is not restricted to a single place in the game world at any given moment in time (Rollings and Adams 2003). Depending on the game, the avatar may be a 2-D, 3-D, animated or non-animated character. Avatars can either be predesigned by the game designers (such as Super Mario in Super Mario Brothers or Lara Croft in Tomb Raider) or customized by the player (such as in World of Warcraft (see Fig. 5.1)).

Fig. 5.1 World of Warcraft avatar creation screen

The act of customizing an avatar creates a feeling of self-expression which can increase engagement (Kim 2009). Players may have the option of choosing their avatar's name/handle, gender, race, class, physical appearance, and other characteristics. When avatars are used in persistent worlds (i.e. a permanent online game world where the state of the avatar continues from session to session), Rollings and Adams (2003) emphasize "the first thing a player does in joining a persistent world is create an avatar, or character who represents her in the game, so this is one of the most expressive things she can do. The avatar is a mask she's going to wear throughout the game" (p.522). Having the control to express oneself in the game can increase the player's engagement (see Sec. 5.4.1 for more details on the benefits of including player control). This personal expression may lead players to have a stronger sense of connection or ownership to their customized characters as opposed to predesigned characters. It can also increase player's physiological arousal (Lim and Reeves 2009). This arousal can lead the player to be more interested in the game, remember more about the experience, and invest more time playing (Reeves and Read 2009). Gee (2003) describes this *committed learning principle* as:

> "Learning involves taking on and playing with identities in such a way that the learner has real choices (in developing the virtual identity) and ample opportunity to meditate on the relationship between new identities and old ones. There is a tripartite play of identities as learners relate, and reflect on, their multiple real-world identities, a virtual identity, and a projective identity" (p.222).

The use of an avatar in an online multiplayer game can also increase the social connection players have with each other. For example, in the massively multiplayer online role-playing game *EVE Online* (*EVE Online* 2003), players pilot space crafts throughout the outer space game world. Players can purchase many different ships, however they may only pilot one space craft at a time. In this sense, the space crafts are the players' avatars, since the players explore outer space solely from the perspective of their ships. However, as their ships encounter other ships they may feel a sense of emotional or social removal, as they are identifying the ships solely as ships, and not the people piloting the ships.

To create a deeper human-to-human bond, the designers of EVE online integrated a mandatory avatar customization. Whereby players create static graphical representations of the characters controlling the ships (see Fig. 5.2). Players can customize their avatars' race, historic background (bloodline & ancestry), gender, costume, facial features (i.e. eyebrows, head, eyes, jaw and mouth), hair, makeup, and skin tone, as well as the background image and lighting. It is these customized characters that the players identify with when playing the game. When piloting their ships, players may target other ships and see the avatars associated with them.

5.2 Game Worlds & Fantasy

Many digital games have a virtual environment where the game play occurs. This virtual environment, the virtual characters and entities within that environment, and the interactions within that virtual environment make the *game world*. The game world can include simple 2-D graphics up to highly realistic 3-D graphics. The aesthetic of the game can also be enhanced with sound effects, music, and speech.

Kiili (2005) states that "the promise of educational games is to engage and motivate players through direct experiences with the game world." (Kiili 2005 p.14). The game

Fig. 5.2 EVE Online avatar customization screen

world is able to create engagement through its ability to immerse the player in a new environment that is separate from real life. Rollings and Adams (2003) explain that "when you go inside the game world and temporarily make it your reality, you suspend your disbelief. The better a game supports the illusion, the more thoroughly engrossed you become, and then the more *immersive* we say the game is" (p.58).

The consequences of exploration and experimentation within the game world remain strictly in the game world, and do not have real life ramifications and there is a freedom to explore without real life consequences (Garris, Ahlers, and Driskell 2002). This allows players to be more daring and perform actions they might otherwise be reluctant to attempt. Gee (2003) calls this the "psychosocial moratorium" principle, whereby players are able to take risks in spaces where real world consequences are lowered. Flight simulators are one example where pilots can practice and learn how to fly without any real life consequences to their well-being if they crash.

Not only can the game world provide the means to try actions which players wouldn't risk attempting in real life, but a game world can be created to allow players to perform actions which simply wouldn't be possible in real life. If the fantasy world is appealing to students, they may be more receptive to the educational content. In fact, Wilson et al. (2009) state that "research on fantasy has suggested that users become more interested than with traditional classroom techniques as well as nonimmersive games and are therefore able to learn material more readily than when using paper-based formats" (p.299).

When applying a fantasy context to educational content, the fantasy can be related or not related to the content. When the content and context are related, the fantasy is labeled as endogenous / intrinsic (Garris, Ahlers, and Driskell 2002; Rieber 1996; Wilson et al. 2009). When the content and context are not related, the fantasy is labeled as exogenous / extrinsic. In the online game *Qwerty Warrior 2*, shown in Fig. 5.3, players shoot and kill approaching enemies by typing words correctly. This would be an example of an exogenous / extrinsic fantasy, as typing and shooting are not related. An example of an endogenous / intrinsic fantasy would be a flight simulator, as the learning material and context are aligned. Garris, Ahlers, and Driskell (2002) summarizes Rieber (1996) who noted that "because endogenous fantasies are more closely tied to the learning content, if the fantasy is interesting, the content becomes interesting. Thus, endogenous fantasies are more effective motivational tools" (p.448).

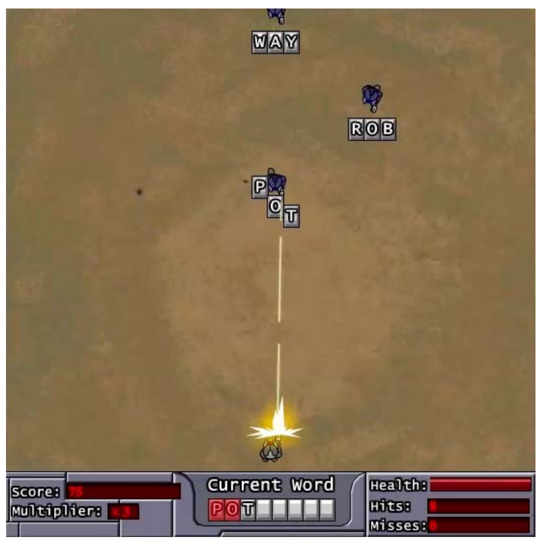

Fig. 5.3 Screen shot of Qwerty Warriors 2 from (McGee 2007)

Another benefit of using endogenous fantasies is that it creates situated learning. Prensky (2001) defines this as "...an approach in which the learning is set in an environment that is similar or identical to where the learning material will be applied in the future" (p.161). Dondlinger (2007) reinforces this concept by stating that "the authentic, situated context affords greater content mastery and transfer of knowledge than a traditional classroom learning" (p.26). The usefulness of situated learning is also noted by Gee (2003) in his *situated meaning principle*:

> "The meanings of signs (words, actions, objects, artifacts, symbols, texts, etc.) are situated in embodied experience. Meanings are not general or decontextualized. Whatever generality meanings come to have is discovered bottom up via embodied experience" (p.224).

Through the use of fantasy in the game world, it is possible to place students in the associated context of the content they are attempting to learn. If it is feasible to deliver the content within its related context, the change for students' to absorb and enjoy the material increases.

A game world with different contexts and virtual landscapes can also assist with memory recall. Reeves and Read (2009) reference a comment regarding the virtual environment of *World of Warcraft* made by James Paul Gee during a presentation to a meeting of investigators in Health Games Research:

> "Every place I pass by again reminds me of a person [avatar] I spoke to here, a challenge I met, a reward received" (p.67).

This concept can be extended to assisting students recall information they gained during a lesson. By being taught a new concept in a new virtual environment, students are more able to compartmentalize this information and recall it later.

Furthermore, having a variety of aesthetics and game mechanics within the game fulfills the player's senses of adventure and exploration. This is especially engaging for players who are *explorers* as described by Bartle (1996). An explorer enjoys the experience of exploring the game world both in terms of its geography as well as its game mechanics.

One game which exemplifies satisfying an explorer's needs is Super Mario Galaxy 2 (Nintendo 2010). In Super Mario Galaxy 2, players are able to play and unlock multiple levels which are selected at the world map (see Fig. 5.4). Each level has its own aesthetic and unique game characteristics. This maintains the novelty of the experience and the exploration. Furthermore, each level is an opportunity to learn new forms of interaction and new aesthetic details (e.g. "This level is bright and colourful, and I can make clouds to stand on... This level is dark and gloomy, and I can shine a light to expose invisible walls."). If each level had the same interactions and aesthetic, the player would become bored and confused as to which level was which. The player would also be less engaged because the monotony of each level would decrease the sense of exploration.

Fig. 5.4 Screen shot of the World Map in Super Mario Galaxy 2 from (Nintendo 2010)

5.3 Story & Narrative

In the same way that it is beneficial to structure the game world to a context which is related to the educational content, the storyline within the game should align with the content. Dondlinger (2007) states "although narrative context does motivate learning, for an educational game to be effective the learning content must align with the narrative plotline" (p.24).

The use of narrative has become an essential element of successful digital games (Rollings and Adams 2003). Kiili (2005) claims that almost every game "has some sort of story attached to it." The use of narrative in games guides the player's actions and organizes character roles, rewards, and group action (Reeves and Read 2009). It sets the background from which all actions are justified, and it gives the player a reason to perform the activities within the construct of the story. For example, in Super Mario Galaxy 2, Super Mario must navigate his way through different levels to rescue the princess from Bowser. This is a common storyline used in the Super Mario video games titles. Without this underlining storyline, Mario is simply running around the game world gathering coins and killing enemies without a purpose. This may be engaging for some individuals. However, adding a story gives the actions a deeper meaning and provides a long term goal for the player.

Conveying information and captivating an audience through storytelling is used in a range of professional fields such as advertising, entertainment, sales, sermons, speeches, case study reports, and journalism (Reeves and Read 2009). Its prevalence in these communication mediums is a testament to its effectiveness. Storytelling conveys information well because content is more easily remembered if it is delivered in a narrative context (Reeves and Read 2009). Reeves and Read (2009) explain that "You can't just give the facts. You need to place them in the context of events sequenced with a beginning, middle and end, some tension about how things will resolve, and detail about the people involved that will engage audiences in something that becomes real because they can imagine themselves in the same narrative space" (p.69).

One way to add tension to a story is by creating mystery and uncertainty. When players do not have all the information, they are left with a mental puzzle to figure out. Salen and Zimmerman (2003) suggest that "...uncertainty is also a narrative concept, for the element of the unknown infuses a game with dramatic tension" (p.388). Mystery occurs when players are aware that there is some information within the game that they do not yet know (Wilson et al. 2009). This creates tension and excitement, which sustains a player's involvement in the game (Reeves and Read 2009).

When players do not know some information, the answer is sometimes provided to them by a character in the game. Non-playable characters can be woven into the storyline in order to offer skill improving challenges to the players. Prensky (2001) refers to this as *coaching*, while Rollings and Adams (2003) refer to the characters as *mentors*. A mentor could be a character who explains at the beginning of the game controls his or her their avatars. In other situations, the mentor can be a character who provides hints and insight for overcoming difficult tasks. Game designers attempt to blend the mentor's lessons seamlessly into the game play, so it does not appear as a learning session (Prensky 2001).

In EVE Online, mentors are represented as *agents* who provide missions to the players. The players can choose whether or not they want to accept the mission. If they do, they will be given further instructions on how to complete the mission. A player's ability to choose makes the video game into a unique narrative medium. While books and movies generally have a single storyline, players can often interact with and alter the storyline in a video game (Gee 2003). The interactivity provided by responsive software ensures that players are not passive observers in the experience. Players become integral in shaping the events of the game, and which makes them more engaged and focused on the activity.

5.4 Flow

When considering ways to create an engaging game, many of the authors listed in Fig. 5.1 emphasize the theory of *flow*. Developed by Mihaly Csikszentmihalyi (1975), flow describes

"...the state in which people are so involved in an activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it even at great cost, for the sheer sake of doing it" (Csikszentmihalyi 1990 p.4). When individuals are in a flow state they achieve:

- High levels of concentration and focus
- A merging of action with awareness
- A loss of self-consciousness
- A distorted sense of time

Csikszentmihalyi refers to this as the *optimal experience*. One example of this would be basketball players who are in the "zone." Their immersion in the game allows for less time to doubt their abilities, which results in a better game. This focus and concentration leads to different perceptions of time, according to (Csikszentmihalyi 1975). Csikszentmihalyi states that the optimal experience is something that we make, and he ascribes specific conditions which can nurture flow. They are:

- A personal sense of control
- Clear goals
- Challenges that are balanced with the individual's skill level
- Immediate feedback

Reeves and Read (2009) reiterate these factors when they state "...people need clear and compelling goals that offer important challenges that match their skills. They must get feedback that is relevant and timely, and they should have an appropriate sense of control over their world. They must experience the satisfaction of meaningful learning" (p.60).

In the next four sections, these attributing factors of flow (control, goals, challenges, and feedback) are examined and discussed with respect to their relationship to digital game-based learning.

5.4.1 Control

"Control refers to the exercise of authority or the ability to regulate, direct, or command something" (Garris, Ahlers, and Driskell 2002 p.451). Within the framework of *flow*, control refers to an individual's sense of control over the outcome of an activity (Csikszentmihalyi 1990). As a player's skill level increases his or her sense of control over the activities also increases (Kiili 2005). For example, a highly experienced dancer will feel more bodily control as she executes difficult dance moves whereas a beginner will feel less control.

If the game is purely based on luck, players will not feel a sense of control. For example, the board game *snakes and ladders* does not offer players a sense of control over the outcome. After players roll the dice, they must move their game piece forward according to the number rolled by the dice. If they land on a game square with a ladder going up, they go up to the indicated square. If they land on a game square that has a snake going down, they go down to the indicated square. Individuals can play the game countless times and their probability of winning will never change. Players are simply following the rules of the game without being able to manipulate the outcome. The only way players could attain a sense of control would be if they learned how to cheat (e.g. controlling the outcome of a dice roll).

Within the context of the proposed electric bass tutoring game, the system must be accurate in extracting and detecting players' performances. If, because of errors in the musical element extraction process, the system detects performance mistakes which players did not in fact make, players will not have a sense of control of the outcome. This lack of control could frustrate players. They could feel cheated and lose faith in the system, and may stop playing the game. This is why the musical element extraction methods described in Ch. 4 must be accurate and reliable.

Control in a digital game can also include the choice over what activities players partake in, or the details about the activities (Dondlinger 2007; Garris, Ahlers, and Driskell 2002; Wilson et al. 2009). In an educational game, the instructional content can be dictated by the software (program control) and/or by the players (learner control) (Garris, Ahlers, and Driskell 2002). In western culture, when learners personally make decisions regarding the type of activities or details about the activities, they have a higher chance of investing more and attempting more complex strategies (Iyengar and Lepper 1999). As discussed in Sec. 5.1, when players can customize aspects of their gaming experience, such as their avatar, they become more invested in the game.

There should be a balance of control between players and the system. If players have too much control they may become bored and unchallenged. If the system has too much control "the player becomes a passive observer rather than an active participant" (Dondlinger 2007 p.25).

A certain level of uncertainty and randomness is necessary in a game. Salen and Zimmerman (2003) explain:

> "Think about it: if you knew who was going to win a game before it started, would you even bother to play?... If a game has no uncertainty - if the outcome of the game is completely predetermined - then any choices a player makes are meaningless, because they do not impact the way the game plays out" (p.174).

For example, when playing the video game *Tetris*, players can control the movement and rotation of a falling block, however they cannot control which type of block will appear next. If players were able to control which blocks appeared there would be no uncertainty within the game. The difficulty of the game would drop significantly, players would not be challenged, and they would lose interest in playing the game.

5.4.2 Rules and Goals

The rules of a game govern the activities of the game. They limit actions and guide players to understand the acceptable and non-acceptable actions within the game world (Prensky 2001). In non-digital games (e.g. board games, card games, sports games) it is the role of one or more human players or referees to determine whether the rules are being upheld. In digital games, the rules are automatically upheld because the computer system or gaming console is monitoring and adjudicating the actions within the game. This allows players to focus solely on the activity of the game, and not on ensuring the rules are not being broken. Reeves and Read (2009) reinforces this concept: "...rules allow players to trust the game. Players value the level playing field created by rules when they're evenly and impersonally applied (after all, it's the computer that most often enforces them)" (p.80).

There are different types of rules that exist in a game. Salen and Zimmerman (2003) offer three categories of rules:

- **Constituative rules**: These are the mathematical fundamentals of the game. They do not explicitly describe how the players should enact these rules. E.g. The rules of physics used in a flight simulation game.

- **Operational rules**: These are the rules that directly and explicitly affect players' behaviours within the game. These are the rules that would be written in the game's instruction booklet E.g. The completion of a flight simulator level occurs when the player successfully lands the plane.

- **Implicit rules**: These are the unstated rules of the game that are implied by the game and assumed by the player. E.g. Players can exit and quit the game when they choose

Salen and Zimmerman (2003) warn "we should emphasize that we do not expect anyone to list every rule of a digital game as it is being designed... [However,] as a game designer, it is extremely important to be able to identify the formal structure of any game you are designing."

The rules of the game are used to impose restrictions and thereby create challenges for the player in completing goals within the game. "In a game, achieving your goals is a big piece of what motivates you... The rules, of course, make this harder, by limiting the strategies at our disposal" (Prensky 2001 p.120).

Goals are useful as people tend to be goal-oriented (Prensky 2001). Goals provide players with a purpose and direction. Clear goals are especially essential, as players will become confused or frustrated if they are not certain what they are to accomplish (Csikszentmihalyi 1990).

Many digital games have different levels of goals: short-term (i.e. in seconds), medium-term (i.e. in minutes), and long-term (i.e. the duration of the game) (Dondlinger 2007). Having concurrent layers of goals can keep players occupied and provide them with satisfaction along their journey of completing the long-term goals.

Players are able to gauge whether they are succeeding or failing to achieve their goals through feedback provided by the game (Garris, Ahlers, and Driskell 2002; Wilson et al. 2009).

5.4.3 Challenge & Game Balance

When designing any game, a designer must consider how easy or difficult it is for players to successfully complete the goals of the game. Players' skill levels must be proportional to the challenges the game offers. Fig. 5.5 shows the relationship between skill and challenge should be related to achieve what Csikszentmihalyi calls the *flow channel*. Gee (2003) refers to this as staying within players' "regimes of competence," whereby the tasks are not too challenging that they cannot be accomplished, and not so easy that players become bored.

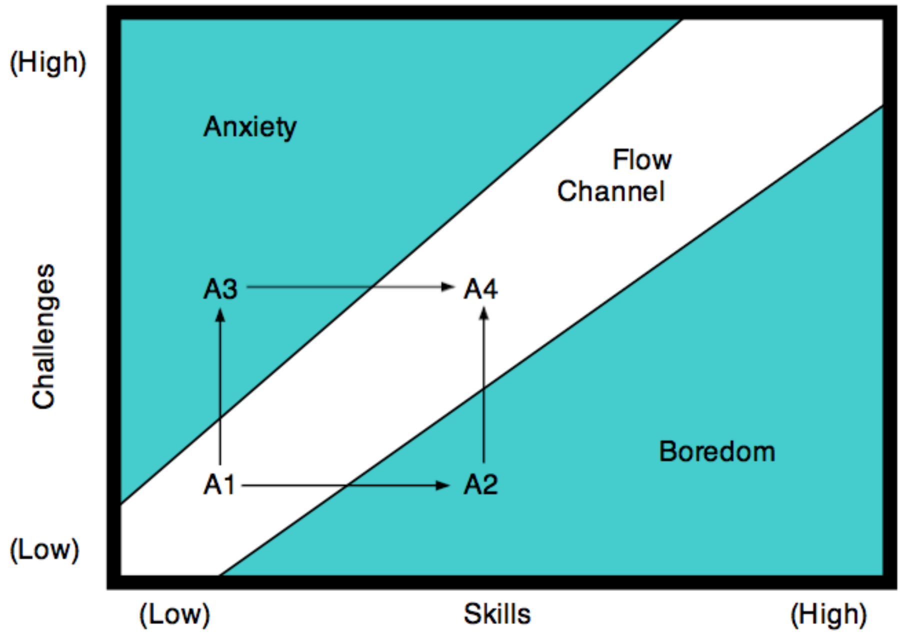

Fig. 5.5 Diagram of relationship between challenges and skills to attain *flow*. Image adopted from (Csikszentmihalyi 1990)

At a beginning state (A1 in Fig. 5.5), players' skills may be adequate enough to handle the challenges facing them. If the challenge is too high (A3) individuals may become anxious or frustrated with the activity. If, instead, players' skill levels increase while the challenge level stays the same (A2), they may deem the activities as boring. Therefore, as the skills of individuals increase the challenge must also increase (A4), or if the challenge increases, students' skill levels must also increase.

However, when designing a digital game to engage players in flow the game cannot control the skill level of the player. It is the responsibility of the game to modify the challenge it offers. In fact, it is one of the strengths of using the computer game medium. Prensky (2001) states:

> "This leads to another important characteristic of computer games - they are *adaptive*. This means that the level of difficulty goes up or down automatically depending on what you do. This is the way computer games keep players in the *flow state*" (p.122).

Piano Tutor, described in Sec. 2.1, provides an applicable example of changing the challenge levels depending on the skills of the players. Although it is not changing the challenge level in real-time, Piano Tutor updates student models to then assign a lesson which better matches a particular student's skill level. If the student's skill level could not handle the previously assigned lesson, Piano Tutor would lower the challenge level by assigning a remedial lesson.

In addition to having the capacity to alter the challenge level to individual players' skill levels, the game itself must be balanced. Rollings and Adams (2003) define "a *balanced game* is one where the main determining factor for the success of the player is the skill level of that player" (p.240). An example of an unbalanced game would be a fighting game where one selectable fighting character is highly overpowered compared to the rest of the fighting characters. No matter which character is chosen, the overpowered character would always win. To resolve this either the other characters could be made more powerful (e.g. they could inflict more damage or have more resistance to receiving damage), the overpowered fighter could be "nerfed" (i.e. the effectiveness of the character could be reduced), or the overpowered fighter could be removed entirely from the game.

Rollings and Adams (2003) state that establishing balance within the game is traditionally a trial and error process, as there are too many variables to be considered (unless the game is fairly simple). Therefore, when developing and implementing the design of a game it may be necessary to test and tweak the balance of the game through trial and error. This is commonly termed *beta testing*.

5.4.4 Feedback & Rewards

In order for players to determine whether they are adhering to the game rules and making progress towards their goals, the system must provide players with an assessment of their performance (Wilson et al. 2009). Traditionally, progress in a game is rewarded by positive feedback and rewards (Kiili 2005).Rollings and Adams (2003) emphasize "for every user action, the game should respond in some way. This one is obvious: When the player interacts with something, he will expect the game to respond - at least with an acknowledgement - immediately" (p.186). Prensky (2001) reiterates "feedback lets us know immediately whether what we have done is positive or negative for us in the game" (p.121).

The immediacy of feedback allows players to instantaneously associate their current actions with the outcome. When the feedback is direct and immediate players can know in that moment whether they should continue what they are doing (or repeat that type of action in the future) or self-correct (or not repeat that action in the future.)

The value of immediate feedback for learning is substantiated through a property of the brain labeled *neuroplasticity*. Neuroplasticity describes the brain's ability to restructure itself through thought and activity (Doidge 2007). One of the core concepts of neuroplasticity is "neurons that fire together, wire together." This describes how the neurons that fire simultaneously (or within a very short time period of each other), increase their associative bond. Doidge (2007) recounts:

> "In 1949 [Canadian behavioural psychologist Donald] Hebb proposed that learning linked neurons in new ways. He proposed that when two neurons fire at the same time repeatedly (or when one fires, causing the other to fire) chemical changes occur in both, so that the two tend to connect more strongly. Hebbs concept was neatly summarized by neuroscientist Carla Shatz: Neurons that fire together, wire together" (p.63).

This explains and emphasizes the importance of offering immediate feedback to players. However, this does not negate the value of offering feedback *over time* (as opposed to real-time). After players complete their missions, it is a common gaming practice to recount the overall accomplishments of the missions. Depending on the game this might entail: the time spent completing the mission, how many coins were accumulated, how many enemies were neutralized, or how many correct notes were played. This gives players an opportunity to reflect upon their overall success at accomplishing their goals.

From a digital game-based learning perspective, this is also an opportunity to review with students what they have learned throughout the mission. This review could include showing quick clips of their actions in critical moments in their game play. This re-exposure to the actions taken during a mission allows players to replay their game playing experience in their mind. Reengaging the neural pathways associated with that experience may improve memory retention and recall.

Feedback can be delivered to players in various forms. It can be numerical (e.g. scoring points), graphical (e.g. in a sports game, seeing your player catch the ball), oral (e.g. other players talking to you), or tactical (e.g. feeling the game controller vibrate as your character is shot) (Prensky 2001). Another invaluable form of delivering feedback is through a game reward or achievement.

When players receive feedback through rewards they are able to quantify the progress they are making in relation to their actions (either in real-time or over time). For example:

the amount of time they have invested in the game can be quantified through experience points (which can be the currency or increase their avatar's "level"); their level of mastery can be quantified by skill points or rank (which increases or decreases depending on their performances); when they are successful at accomplishing defined goals they can receive achievement badges or some virtual currency.

As players increase their experience points, skill levels, or achievements new features of the game can become available. These new features (e.g. new missions, new abilities, or new customizable avatar graphics) can be automatically unlocked by the system, or players can personally choose which features to unlock by redeeming some form of virtual currency.

Having the ability to unlock aspects of the game gives students tangible goals to work towards. It also ensures that the possibilities of game-play are not given away all at once. Instead, players are eased into each new aspect of the game. This is crucial in establishing good game pacing and it taps into the engaging game mechanic of collecting (Kim 2009, 2010).

A game which has points and lets players level-up their avatar lends itself to fulfilling the needs of *achievers* (Bartle 1996). As their namesake implies, achievers thrive on achieving goals and improving their standing within the game. Achievers tend to be competitive and strive to be the best at in-game tasks. Multiplayer games can use a leaderboard to highlight the achievements of the top players. Having the means to quantify and broadcast their accomplishments through a leaderboard is an effective way of satisfying the needs of achievers.

5.5 Social Interaction

More and more digital games are designed to be multiplayer . Reeves and Read (2009) indicate "there has been a sea change in the popularity of solo games compared with those that involve multiple players. Group games, whether played at a computer, a console, or even casually on a phone, are winning" (p.82). One reason group games are becoming more prevalent, compared to solo games, is they offer social interaction.

Although game programmers attempt to create realistic artificial intelligent (AI) characters, AI cannot provide the same depth of interaction that human players can offer (Prensky 2001). Knowing that humans are controlling the actions of the characters they are interacting with in the game increases the social authenticity. When players interact with other human players their primal social instincts are engaged and fulfilled. Reeves and

Read (2009) explain “the close connection to other players causes one’s social-emotional engine to run continuously, and that charges interactions” (p.82). The social interaction with digital games can consist of social communication as well as cooperative or competitive play.

There are a variety of ways to communicate with other players within a game. Players can communicate directly through:

- real-time text, voice or video chat
- in-game e-mail messages
- discussion forums / message boards
- posting text or video comments
- rating user content

These communication channels allow players to articulate and convey their ideas to other players. Having communication channels is necessary to engage players who are described by Bartle (1996) as *socializers*. Socializers enjoy the social interactions and inter-player relationships. Bartle (1996) states this can include “empathising with people, sympathising, joking, entertaining, listening; even merely observing people play can be rewarding - seeing them grow as individuals, maturing over time.”

Details of players’ histories and expertise can be communicated to other players through their avatars or personal profile pages. Players’ levels, ranks, achievements, acquired items, and other details can be publicly viewed by other players. These details can be used as a shorthand to determine players’ proficiency within the game. A highly ranked player with numerous rare items can quickly be viewed as a player with high proficiency within the game. This taps into the same psychological stereotype shortcuts people may apply in real life, such as a well groomed man dressed in a fancy suit with an expensive watch probably has a good monetary income (though this may not necessarily be the reality of the situation).

Players can also communicate with others through the movements or actions of their characters within the game. The “body language” of their character can indicate their intentions to other players. For example, in a competitive first player shooter game, after a player has killed another player he can position his character over the dead body of the defeated player. This is a form of taunting which the defeated player will see on his screen until his slain character respawns.

Competitive play within a digital game gives players an avenue to put their skills to the test against other players. Competition challenges players to apply and improve their skills. Competitors can range from those having well-intentioned sportsmanlike attitudes to those with ill-mannered cut-throat temperaments. The latter would represent the players described by Bartle (1996) as *killers*. Killers enjoy imposing themselves on the other players. Killers' entertainment comes from inflicting distress upon other players. In an educational setting it is common practice to ensure that other students are not inflicting distress upon others. However, if students who are intrinsically motivated by imposing themselves on others are squelched entirely they may lose motivation for doing the activity. An appropriate investigation would be necessary to determine how much leniency could be offered to *killers* in the context of educational games. For example, boasting and taunting other players may not be encouraged however these actions could be minimally tolerated if they did not completely disrupt or hamper the learning potential of other students.

In cooperative play, on the other hand, players are able to work together towards a common goal. Within the game, players perform actions which are mutually beneficial towards this goal. Sharing information can also be a cooperative activity which increases engagement. Salen and Zimmerman (2003) state "online communities can also be used to share information to solve problems in the game. Games can be designed to encourage players to interact socially. This collaboration can heighten the satisfaction of the game play" (p.576).

When players share information and work together within a game they create *community of practice*. Gee (2003) labels this an *affinity group* where learners form a "group that is bonded primarily through shared endeavours, goals, and practices and not shared race, gender, nation, ethnicity, or culture" (p.227). When members of the affinity group are able to work collaboratively and share information with their more experienced peers they are able to learn within the"zone of proximal development" as defined by Lev Vgotsky (Gee 2003). This can improve the students' capacities for learning.

5.6 Summary

This chapter reviewed the research of experts in the field of education, digital game-based learning, and video game design. From their works, a list of digital game elements was synthesized. The notable results include:

- Allowing players to use customizable avatars increases control and self expression
- Integrating game worlds and fantasy immerses players within the game
- Setting the game in a context which is related to the content leverages situated learning
- Using a narrative plotline that aligns with the learning content adds deeper purpose to players' actions within the game
- Engaging players in *flow* fosters high levels of concentration and focus, a merging of action with awareness, and a loss of self-consciousness
- Facilitating a balance between player and program control ensures players do not become bored or passive learners
- Establishing clear rules and goals guides players through the learning activities with purpose and direction
- Offering a level of challenge that matches players' skill levels keeps players in a state of *flow*
- Providing players with immediate feedback allows them to instantaneously associate their current actions with the outcome
- Creating a framework for social interaction among players engages players' social-emotional engines

Chapter 6

Proposed Game Design

The following sections describe the high-level requirements of the proposed electric bass tutoring game. As this is a high-level design, the overall concepts will be outlined while their exact implementations may not necessarily be specified (e.g. I will state that the game will allow the player to customize an avatar, however I will not specify the aesthetic style which should be used to depict the avatar). The specific ways this design is implemented would need to be outlined in a separate document describing an implementation strategy. At times, I may describe possible ways of implementing a game requirement, however these implementation examples strictly given to convey the concept of the requirement.

Each game requirement is based on the principles explored and defined in the previous chapters. Throughout this chapter, I will state chapter and section numbers for any related concepts for cross referencing purposes.

One of the foundational concepts of the proposed game is the theory of situated learning (Sec. 5.2). The main premise of the proposed game is the players are up and coming bass players. Many of the activities in this game are based on activities real professional bass players may encounter sometime throughout their careers. The players are able to practice these activities in a virtual setting without risking any real life ramifications (e.g. embarrassing poor performances, losing money on a gig, etc.). Performing these activities in the game offers players experiential learning. After playing the activities in a virtual setting players may become more prepared and comfortable performing these activities in real life.

Each game requirement is categorized as one of the five components:

- *Game Feature*: visual, auditory, or conceptual parts of the digital game (e.g. the game world in which the game takes place)

- *Bass Activity*: game activities in which players physically play their electric basses (e.g. playing a virtual gig)
- *Social Activity*: game activities in which players interact with other human players (e.g. customizing the avatar)
- *System Process*: the tutoring system processes and/or presents information within the game (e.g. determining and displaying feedback of players' performances)
- *Database*: digital storage of game related content (e.g. a repertoire database stores all the music which can be played by the players within the game)

Fig. 6.1 contains an overview of the requirements for the proposed electric bass tutoring system. In the game, players' *avatars* are guided within the *game world map* by the game's *storyline* to encounter a variety of *non-player characters* (NPCs). The NPCs present the players with *bass activities* to be completed using their electric bass. Different *system processes* are used to successfully facilitate and offer feedback before, during, and after the bass activities. A few *databases* are used to store pertinent game related data. Players are also able to engage in a number of *social activities* with other human players.

Fig. 6.1 Overview of the proposed electric bass tutoring system

This chapter will describe the details for each game requirement.

6.1 Game Features

6.1.1 Avatar

Players go through the game from the perspective of their avatars (Sec. 5.1). From the beginning of the game, players customize the physical characteristics of their avatars (e.g. gender, head shape, hair style, hair colour, eyes, eye brows, mouth, nose, ears, body type, height). The physical appearance can also be customized through clothing and accessories (e.g. hats, shirts, pants, shoes, earrings, tattoos). New customization items (e.g. hair styles, shirts, pants, etc.) can be unlocked or purchased later in the game (Sec. 6.5.3. Players complete the avatar customization process by giving it a name.

After the initial avatar customization process, players will also specify how long they have been playing electric bass in real life and what skill level they believe they are between a range of 1-10. This will help the system initially set the challenge level to match their estimated skill level (Sec. 5.4.3).

Players will also indicate which style of music they enjoy. This may include selecting bands they enjoy listening to or playing. This information can be useful, as it will be easier to engage players if they are learning musical pieces from bands they already have an affinity towards.

6.1.2 Student Model (Database)

When players first enter the game the system does not know the players' actual level of musical proficiency, but as players progress the system will track and gauge their performance levels. The history tracking of the players' performances will be stored in the players' *student models* (Sec. 2.1). These student models will help the system determine whether students have the appropriate skill levels to succeed at an exercise. The system could then suggest exercises that are challenging enough to keep players in the flow channel (Sec. 5.4).

The student model has a record of the players' activity histories as well as any *achievements* they have earned. Achievements are awards given to players in the form of a graphical badge when they have completed a task. This is similar to the badges boy scouts receive for completing specific scout related tasks. Achievements have been used in Xbox Live, Star Craft 2, and in the casual online game portal www.kongregate.com. These achievements give players goals to work towards as well as quantify their accomplishments (Sec. 5.4.2). An example of an achievement could be a "first timer" achievement after players have successfully completed their first gigs. The achievements could also include real life examples such as the musical achievements of performing new gigs, making CDs, playing with a new musical group, or musical awards such as grammies.

The student models will also retain player stats (a.k.a statistics) which are represented by different point systems. The points systems created for this game include:

- *Playing Experience Points (PXP)*: this increases when players engage in activities
- *Level*: a shorthand of the PXP. E.g. 10 PXP = level 1; 1000 PXP = level 10
- *Musicianship Points (MP)*: this measurement of skill increases or decreases depending on whether players are successful at completing activities
- *Rank*: a shorthand of the musicianship. E.g. 10 MP = beginner; 1,000 MP = novice
- *Reputation Points (RP)*: this measures the amount of respect and trust non-player characters[1] (NPC) have for the player. This represents the reputation players have in terms of their professionalism and musical capabilities (e.g. players' RP will decrease if they do not attend a scheduled gig). Each NPC has a unique amount of RP related to the players.
- *Regional Reputation Points (RRP)*: the overall musical reputation in a region, which is calculated from the sum of RP from all NPCs in that region
- *Fans*: this is the number of fans the player has within specific regions (i.e. within neighbourhoods, cities, provinces, and countries)
- *Funds*: virtual currency used in the game

The way these stats are changed throughout the game is shown in Sec. 6.2.

Similar to achievements, the system uses these stats to determine what activities will be offered to players. For example, when players have reached a certain *rank* the ability to do a new activity, such as a recording session (Sec. 6.4.5), could be unlocked.

[1] Non-player characters in this sense are characters within a digital game which are not controlled by a human player. This is not to be confused with the term *non-player* which may be used to refer to individuals who do not play music or a specific instrument.

6.1.3 Game World Map

In the proposed game, players can navigate to different selectable activities through a game world map similar to the one used in Super Mario Galaxy 2 (see Fig. 5.4 in Sec. 5.2). However, in this game players navigate the game by selectable cities and then neighbourhoods within those cities. Each neighbourhood will have different graphical buildings which represent activities which players can select and attempt. For example, a neighbourhood could include players' houses where they practice (Sec. 6.4.2), music studios where they take lessons (Sec. 6.4.3), and performance halls where they perform gigs (Sec. 6.4.4).

Having a graphical representation of the location for each activity will help players compartmentalize the activities in their minds (Sec. 5.2). If they reflect upon a lesson they took, it may be easier for them to recall that information simply by thinking about which building the lesson took place in.

A 3-D virtual environment, such as the one used in World of Warcraft, can be immersive but it would be unnecessary for this application. 3-D virtual environments offer players the ability to interact with virtual objects and learn from that spatial experience. This can be useful for flight simulation games. However, within this game the players' main spatial interaction is with the real electric bass in their hands. Furthermore, navigating their characters through an expansive 3-D virtual environment would take up too much of their music practice time. The time spent on the game should be optimized for an engaging experience related to the musical activities within the game. If players want to play an activity they shouldn't have to waste time travelling to it.

New buildings/activities, neighbourhoods, and cities will become unlocked and available from the world map as players progress through the game.

In the proposed game the main buildings are: the players' houses, mentors' teaching studios, rehearsal spaces, recording studios, concert halls and bars. The activities related to these locations are discussed in Sec. 6.2.

6.1.4 Narrative & Storyline

A storyline guides players through the game. The use of narrative will offer players a sense of immersion into the game world and deepen the justification of their actions within the game (Sec. 5.3).

This thesis will not outline all the details of the story (e.g. character names, plot twists, conflicts) which should be implemented, as that would require research into narrative studies. However, for the sake of having an example which fits with the learning electric bassist, I will borrow the story used in Victor Wooten's book *The Music Lesson: A Spiritual Search for Growth Through Music* (Wooten 2008a).

In this story a young bass player encounters a music teacher. The music teacher teaches the main character about the different musical elements (Ch. 3). The music teacher helps the young bassist understand musical concepts by giving analogies to nature. Throughout the book the young bass player also encounters different teachers who give him deeper insights into music.

In the proposed game, the delivery of the musical educational content does not necessarily need to relate music to nature. However, the main storyline of the proposed game will revolve around players' bass playing avatars who encounter different characters who teach them about music and the electric bass. Each character will interact with players based on a specific type of bass activity.

6.1.5 Non-Player Characters

Throughout the story, players meet a variety of non-player characters (NPCs). As described in Sec. 5.3, NPCs can be used within a digital game to teach players new concepts. In the proposed digital game, the NPCs will act as the teachers delivering the educational content, giving players exercises to perform, and offering feedback regarding their performances. The benefit of having a graphical character deliver the educational content instead of simply showing instructional text is it puts a *face* to the source of the educational content (Sec. 5.1).

Within the game there are five specific types of characters. There are *band members*, *bookers*, *mentors*, *recording engineers*, and *sound guys / girls*. Each type of NPC is used as the *face* to communicate information to the student within the context of specific activities within the game (Sec. 6.2).

- *Band members* play gigs and practice music with players
- *Bookers* book players to perform at their music venues
- *Mentors* offer players music lessons
- *Recording engineers* work with players during recording sessions
- *Sound guys/girls* ensure the system can properly detect the players' bass performances

The NPCs communicate with the player through speech, text, and graphical animation or video. The NPCs could be 2-D or 3-D computer animated characters or live video / audio recordings of real humans. Voice actors should be used and recorded to accommodate speech, while voice synthesis should be avoided. Most successful digital games do not use voice synthesis for their characters, unless the character is supposed to be a computer. Players can pick up on the inauthenticity of voice synthesized speech and it detracts from their immersion in the fantasy of the game world. Most games either use voice acting or strictly use text.

In the same way that different locations in the game world can be used for different activities (Sec. 6.2), different NPCs can be used for different learning materials. For example, one mentor could focus on teaching reading music (e.g. identifying the symbols used for the different note duration values), another mentor could focus on playing techniques (e.g. slapping and popping), and another mentor could focus specifically on a genre of music (e.g. reggae). The content of the lesson would depend on the implementation strategy.

Another benefit of using a variety of mentors throughout the game is that it allows the game to offer different personalities and teaching styles. Every student is unique, therefore if students do not enjoy the personality or teaching style of one teacher, offering them the choice of another may increase their engagement and overall learning.

Each NPC will have a certain level of RP points associated with the players. The higher the RP, the more the NPC will offer players opportunities or refer players to other NPCs (which will unlock more activities). If the RP points become too low, that NPC may stop associating with the player and any activities related to that NPC will be removed. The RP will rise or fall depending on players' successes or failures at performing the bass activities.

6.2 Overview of Bass Activities

There are eight types of bass activities players may engage in throughout the game:

- *Sound and Video Check*: players check that they are in tune, their dynamic levels are set appropriately, and the computer vision system is working correctly (Sec. 6.4.1)
- *Woodshedding*: players practice whatever they wish using performance tools (Sec. 6.4.2)
- *Lessons*: players receive private lessons from mentor NPCs (Sec. 6.4.3)
- *Gigs*: players perform virtual live concerts (Sec. 6.4.4)

- *Recording Session*: players perform musical pieces in the context of recording sessions (Sec. 6.4.5)
- *Auditioning*: players attempt to try out for a band (Sec. 6.4.6)
- *Rehearsing*: players practice playing musical pieces with a virtual band (Sec. 6.4.7)
- *Lifting*: players attempt to figure out the bass line of a musical piece by ear (i.e. listening and mimicking the bass line)

These bass activities are described in more detail in Sec. 6.4. The mechanics of the bass activities are all fairly similar. The main structural framework is adapted from the lesson structure used by Piano Tutor (see Fig. 2.2 in Sec. 2.1) and VEMUS (Sec. 2.4). The individual structure used for each activity will involve slight alterations to the structure described below:

- A musical piece (Sec. 6.6.3) or lesson topic (Sec. 6.6.2) is selected (either by the system or students)
- The NPC presents the content of the lesson (Sec. 6.3.1)
- The NPC explains the exercise by stating clear goals (Sec. 5.4.2)
- Players select, if they wish, to use any of the *performance tools* (Sec. 6.3.3)
- Students perform the exercises (Sec. 6.3.2)
- The audio, MIDI and video data of students' performances are recorded (Sec. 6.3.5)
- The NPC and system provides real-time feedback (Sec. 6.3.4)
- When the exercise is complete the system performs an offline analysis of players' performances (Sec. 6.3.4)
- The NPC offers feedback to students regarding their overall performance (Sec. 6.3.4)
- The NPC indicates whether the exercise was completed satisfactorily
- Players' student models are updated (Sec. 6.1.2)
- Players can review the audio and video of their performances (Sec. 6.3.5)

- Players can choose whether to save or delete the recorded performances (Sec. 6.3.5)
- Players can also choose to share the recorded performances (Sec. 6.5.4)

The methods of musical extraction, analysis, and feedback are taken from the findings from Ch. 4. Sec. 6.3 outlines the system processing which occurs for the bass activities.

After the bass activity is completed players can optionally rate the bass activity or the NPC delivering the content. This helps create a social exchange between players as they can see which bass activities and NPCs are popular and may be worth attempting.

If players succeed in completing the goals of the exercises they will be rewarded with improvements to their stats and/or with unlocking aspects of the game (e.g. unlocking activities, world map regions, achievements). If players do not succeed in completing the goals of the exercises they may or may not be reprimanded by a decrease to the players' stats. The amount of increase or decrease to the players' stats is dependent upon the type of activity.

6.3 System Process

6.3.1 Deliver Content & Instructions

The NPC presents new music subject matter and offers instructions for the related bass activity. The presentation of the material may include multiple sensory form such as audio clips, music clips, speech, graphical images, animated videos, and / or videos demonstrations. The use of multiple sensory modalities can improve the means of communicating information to students. The mentor will also specify clear goals regarding the related exercise (Sec. 5.4.2).

The performance tools also give players the ability to select and loop specific sections during the practice songs and woodshedding activities.

6.3.2 Exercises & Real-Time Feedback

When performing bass exercises, players will have a graphical representation of what they are to play on their basses. This could be represented with musical scores or tab notations similar to the method implemented by the Interactive Software for Guitar Learning (Sec. 2.8).

As players play the bass, the system will follow the performance using score following and will provide real-time feedback regarding the execution of the musical elements

(Sec. 5.4.4). Real-time feedback could be implemented using a colour coding representation similar to the one described in Sec. 2.8. Depending on the exercise, the system could make use of the "wait mode" described in Sec. 2.8.

A graphical representation of the fretboard will be used to help demonstrate to students where their fingers should be placed on the bass (Sec. 2.5, Sec. 2.3, & 2.4) Feedback can be offered to players orally through mentor comments. For example, the mentor could say "Great job!" or "A little off there!" after the bassists perform a challenging phrase.

Feedback could also be given by points which are awarded when players play aspects of the exercise correctly (e.g. the correct pitch, rhythm, articulation, dynamics, phrases, etc.).

6.3.3 Performance Tools

During a bass activity, players are able to use a virtual metronome, virtual tuner, virtual fretboard (which indicates which strings and frets are to be played), accompaniment, score / piano roll / tabulator musical representation, a tempo adjuster, and graphic curves related to players' performance (Sec. 2.8 & Sec. 2.4). Players are able to customize which performance tools will be enabled / disabled during the bass exercise.

6.3.4 Performance Analysis & Post Performance Feedback

When players complete an exercise the system analyzes the overall performance. The feedback offered by the system will focus on which musical elements are listed as most pertinent to that particular exercise. The system will determine the top 3 most pertinent feedback messages (one of which will always be positive). These 3 feedback messages will be conveyed by the mentor to students through text, speech, mentor animation / video, as well as by playing back the audio and video of the students' performances on which the 3 feedback messages were based.

The system will also generate any graphic curves which may assist players in listening to or understanding their performances (Sec. 2.4).

The system will determine whether performances meet the necessary requirements to consider the exercises a pass or a fail. The mentor will then convey this information to the players. The players' student models will also be updated.

6.3.5 Record and Review Performances

During bass exercises, the audio, MIDI, and video of students' performances are recorded. Since the digital storage size required to save the audio, MIDI and video may be large, after the bass exercises are complete the students may choose to save or discard the recorded performances. Reviewing audio, MIDI and video recordings of students' performances will help students self-assess their own performances. Furthermore, these recordings will archive an auditory and visual history of students' progress.

6.4 Bass Activities

6.4.1 Sound and Video Check

The game will use a real bass and incorporate the hardware requirements described in Sec. 4.5. The game will need to ensure that all the hardware is set up appropriately. This is done through a *Video and Sound Check* activity. Using a virtual tuner provided by the system, players will verify that each string is in tune.

The system will also determine whether players need to adjust any volume settings to ensure the system can properly detect loud and quiet dynamics. Finally, the system will verify that the web cams and ARToolKit are functioning correctly.

Once the sound and video check is complete players can proceed.

Players will be required to quickly do a tuning and dynamics sound check before they perform any activity that involves using their basses. This will improve the reliability of the pitch, dynamic, and computer vision information received by the system. Because there is a divided pickup under each string, players can verify tuning simply by strumming the strings. If each string is in tune, the system will acknowledge this and allow players to proceed. In one possible implementation, a graphical green light and a pleasant ding sound could provide feedback to students regarding a tuned string. A pleasing aesthetic can make this activity a rewarding experience and provide positive reinforcement (Sec. 5.4.4).

6.4.2 Woodshedding

Woodshedding is a slang term referring to practicing music. When players perform the activity of woodshedding there is no explicit lesson or goal. The student's are able to use the Performance Tools to explore their own performance. The system can record players' performances for reviewing purposes. Players' PXP will increase depending on how much time is spent practicing.

6.4.3 Lesson

One of the main activities in the game is taking music lessons. Bass lessons can be selected from the mentor's music studio shown on the game world map. For each lesson, bass players will play an exercise to demonstrate their mastery of the lesson content. The content of a lesson is either predicated by the system (using the players' student models) or chosen by the players. The structure of the lesson follows the structure listed in Sec. 6.2. The exercise may be a musical pattern, a musical piece, or a portion of a musical piece.

It is possible for a lesson to focus solely on one or more musical elements. This will allow players to focus their attention on the execution and understanding of that/those musical element(s). For example, an exercise may be strictly a rhythmic exercise. Therefore, playing specific pitches is not relevant to the exercise and should not affect the grading of students' performances. The mentor will clearly specify which musical elements are to be considered for the exercise, as clear goals is a key factor in creating flow (Sec. 5.4). If players are successful in completing the activity their PXP, MP, and RP will increase. If players are not successful only their PXP increases.

6.4.4 Gig

As the storyline unfolds, players will be offered gigs (i.e. a job performing live music). A large part of getting gigs is having a good RP with venue bookers or to have high RRP within a specific region (i.e. neighbourhood, city, province, country).

When offered a gig, players can choose to accept, decline, or reschedule it. The terms of the gig will be specified to help players decide this. The terms may include how their stats will be affected if they perform well for the gig (e.g. PXP +10, MP + 10, RP +10, fans +10, funds +100).

The structure of performing a gig in the game is similar to the structure of a lesson, however, there are some differences. One of the main differences is that players are not able to stop the song and try again. When players perform a musical piece during a gig activity they only have one shot at it. If the gig doesn't go well players will have to wait until they are offered another gig before they get a chance at redeeming themselves.

Depending on the gig, a musical score / piano roll / tabulator representation of the musical piece may or may not be graphically shown to players. The reason for this is that in real life there are gigs where players must perform *by heart* (i.e. perform without the aid of reading a written musical representation). For example, it would be unusual if the members of Metallica performed to their heavy metal audience while reading sheet music on music stands.

For a gig, players may be requested to play more than one musical piece in succession. This is the structure of the gig activity:

- One or more songs are selected by the system or players from the players' repertoire (Sec. 6.6.3)
- Students perform the song(s) along with musical accompaniment
- The audio, MIDI and video data of students' performances are recorded
- The system / crowd provides real-time feedback (Sec. 6.3.4)
- The system performs offline analysis of players' performances when the exercise is complete
- The system indicates the crowd's response to the overall performance
- The system indicates whether the gig was a success or failure
- Players' student models are updated
- Players can review the audio and video of their performances (Sec. 6.3.5)
- Players can choose whether to save or delete the recorded performance
- Players can also choose to share the recorded performance (Sec. 6.5.4)

If players play a successful gig their PXP, MP, RP, fans, and funds will increase.

If players do not perform well for the gig their PXP will increase, their funds may or may not increase (depending on the type of gig), and their MP, RP, and fans will decrease. As in real life, after a poor live performance, the income generated from merchandise or CDs sold to fans will decrease, and respect towards the musician will decrease both within the musical community and from the fans.

Another aspect of the gig activity which is different from the other activities is that some gigs must be played on specific days. The passing of time matches the passing of time in real life. In real life, working bassists are booked to play gigs for specific dates and times. Being at the gig location at the right date and time is part of being a reliable musician.

In the game, players will book gigs to be played on real life days and time. To fit in with the players' real life schedules, players will be able to accept, decline, or reschedule the date / time of these gigs.

Not only will this simulate the real experience of being working bass players, it is also a reason for players to return to the game. This can sustain players' engagement by setting goals for the future. Players will want to work towards the gig that is coming up in a few days / hours.

If players do not attend their gigs their musical reputations will decrease (Sec. 6.5.2 for details on musical reputation). It will especially decrease with the NPC within the game who got them the gigs, as well as the venue. However, players have the open to find a substitute player to take the gig for them. See Sec. 6.5.1 for more details on "subbing."

6.4.5 Recording Session

Recording sessions are available from recording studios displayed on the game world map. A recording session is similar to a gig except players are able to perform multiple *takes* (i.e. they may retry playing the performance). Furthermore, if players play the majority of their performances correctly, they can specifically retry the trouble sections.

Instead of interacting with a mentor, players interact with a sound engineer character. The engineer communicates with players regarding whether they performed the music correctly or not. However, he does not necessarily provide insight about ways to improve.

If the recording session is successful players' PXP, MP, RP, fans and funds increases. If the recording session is not successful players' PXP increase, however their RP decreases.

6.4.6 Auditioning

Auditioning is similar to performing a recording session, however, the context and outcome of the activity is different. Auditions take place at rehearsal studios located on the game world map.

If the audition goes well players' PXP, MP, and RP will increase. Players will also be asked to perform gigs with this band in the future. If the audition does not go well players' PXP increase while their RP may decrease and this band will not offer future gigs to the players.

6.4.7 Rehearsing

Rehearsing is similar to auditioning (see Sec. 6.4.6), however the purpose of the activity is to repeatedly practice musical pieces with a virtual band. Rehearsals take place at rehearsal studios located on the game world map.

If the rehearsal goes well the players' PXP, MP, and RP increases. If the rehearsal goes very well, players may be given the opportunity to reschedule an upcoming gig to an earlier date. If the rehearsal does not go well the players' PXP will increase, however their RP may decrease.

6.4.8 Lifting

Lifting a musical piece refers to figuring out the musical part of one or more instruments by ear. This involves listening and imitating the musical performance.

Within the game, lifting is similar to practicing a song, however the musical notation is not graphically shown until players successfully play passages of the song. Players attempts to play the correct bass line while listening to the musical piece (both the accompaniment and the bass part). Players will also be able to loop and/or slow down sections.

Players must play a phrase (though not necessarily the phrasing) of notes correctly to unlock the sections. This ensures that players do not randomly play notes with the hope of haphazardly piecing together the musical piece by chance.

From the game world map, players can select the lifting activities from a building representing their homes. Players can purchase *liftable* songs with funds from a graphical representation of a music store. This simulates the concept of purchasing music and returning home to learn how to play it.

If a song has been successfully lifted players' PXP and MP will increase and the music piece is added to their repertoire (Sec. 6.6.3). If players cannot lift the musical piece, they are still rewarded for trying by increasing the players PXP.

6.5 Social Activities

6.5.1 "Subbing"

Since players may not always be able to schedule playing a gig they agreed to, there is also the option of finding another game player to take their gig. In real life, this is what some musicians would do if they found out they couldn't attend the gig. This is sometimes referred to as "subbing in," as the other player is substituting for the the original musician.

In the game, if another player agrees to take the gig, it will be removed from the initial player's calendar and added to the subbing player's calendar. The RP for substituting players will increase, while the RP of initial players will go down slightly but not as drastically compared to if they did not attend the gig without finding a subbing player.

6.5.2 Player Profile Page

Similar to social media sites, players' profile pages display details about the players' stats, history, accomplishments and other details. Players can customize which data is made public to others. Types of content on players' profile pages include:

- Image of players' avatars and names
- Stats
  - Playing experience points (PXP)
  - Level
  - Musicianship points (MP)
  - Rank
  - Reputation points (RP)
  - Regional musical reputation (RRP) (with a method of selecting which region)
  - Fans
  - Funds
- Calendar (e.g. upcoming / previous gigs, recording sessions, etc.)
- Badges and achievements
- Gear
- Social Media
  - Friends (human players)
  - Shared videos

  - Favourite lessons
  - Favourite NPCs

- History
  - Number of gigs played / missed
  - Number of lessons taken
  - Number of musical pieces recorded from recording sessions
  - Number of songs in repertoire
  - List of affiliated NPC (e.g. venue bookers, band members)

6.5.3 Redeem Virtual Currency

Players can spend their virtual currency on virtual items for their avatars. Items for avatars include appearance modifiers (e.g. hair style, clothing, facial features) and musical gear (e.g. basses, bass amplifiers). These items are solely for display purposes and they can be bought and sold to other players. This can engage players who enjoy acquiring achievements in digital games (Sec. 5.4.4).

6.5.4 Performance Sharing

When a bass activities are completed, players can review the audio and video recording of the performances. Players can then choose whether they wish to share these performances with others. Players can share the audio, video, as well as what was graphically seen by the players (i.e. the user interface and any real-time system feedback). When players select bass activities they can view any shared performances related to those activities. This can help players learn from the experience of others (Sec. 5.5).

This can be a means for players to brag to others about their accomplishments. It can also be a way for players to share their opinions and methods for succeeding at different activities.

6.5.5 Message Board

Within the game players can post comments and have open discussions regarding the game or other topics that are related to music and the electric bass.

This is a channel for players to ask questions, provide answers, and share ideas with the other players. They may also post videos of their performances and ask for feedback (Sec. 6.5.4). This is similar to individuals who might post performance videos on youtube.com, however, the individuals watching the performance videos will also see the feedback that the system is offering. This is effective at engaging players that enjoy socializing. It can also create a framework that encourages a community of practice (Sec. 5.5).

6.5.6 Leader Board

A leaderboard can be used to encourage competition among players (Sec. 5.5). High scores, player stats, achievements, highest RP, highest PXP, etc. can be compared across the players and the top players can be displayed. Daily, weekly, monthly, and yearly top players can be viewed. This will engage players who enjoy competition, and it will allow players to see how they measure up against the other players.

6.6 Databases

6.6.1 NPC Database

All the data related to a NPC is stored in the NPC database. Each NPC delivers content to players (Sec. 6.3.1) and offers them feedback (Sec. 6.3.4). The audio, videos, and graphics related to delivering a lesson and providing feedback are stored here. For example, feedback messages from the mentor (e.g. "Great Job!" or "The time of that note is a little off.") are stored here.

6.6.2 Activity Database

The activity database contains all the necessary data to deliver a bass activity to students, including: related NPC, delivery content (e.g. videos, audio files), a reference to the exercise or musical piece to be played by the players, and any specific feedback messages that are specific to this activity.

6.6.3 Repertoire Database

In order for the system to know whether students have performed a piece music correctly, it must have a pre-existing representation of what it expects from the performances. All of the musical pieces and exercises are stored in the *repertoire* database. It should be noted that legal issues would need to be addressed for obtaining the rights to implement and use any songs which have copyrights attached to them.

When playing the game, players will not immediately have access to all the musical pieces and exercises in the repertoire database, as this may overwhelm them. Instead, they are unlocked throughout the course of the game by completing activites. A running tally of all the songs players have unlocked are documented within the game as their own personal *repertoire*. This is similar to real life musicians who may have a collection of musical pieces they know and have in their repertoire.

6.7 Implementation Limitations

Although there are no technological limitations to implementing the proposed design (as all the concepts in the game have been previously implemented in other sources), there may be some economic limitations.

The cost of developing and testing a software of this magnitude would be expensive. Therefore, an implementation of the design may require removing many of the requirements.

Implementing a *bare bones* version of the design may be a sound alternative. Additional requirements could be introduced later.

6.8 Summary

This chapter outlined the proposed design of an electric bass tutoring game. Table 6.1 summarizes the the relationship between the bass activities, locations in the game world, and the related NPC.

Bass Activity	**Location**	**Related NPC(s)**
Audition	Rehearsal Space	Band Members
Gig	Concert Hall / Bar	Booker & Band Member
Lesson	Mentor's Studio	Mentor
Lifting	Player's House	(just Avatar)
Recording Session	Recording Studio	Recording Engineer
Rehearsal	Rehearsal Space	Band Members
Sound and Video Check	NA	Sound Guy / Girl
Woodshedding	Player's House	(just Avatar)

Table 6.1 List of Bass Activities

Table 6.2 displays how succeeding or failing to complete a bass activity affects the players' stats.

Bass Activity	**PXP**	**MP**	**RP**	**Fans**	**Funds**
Audition	▲/▲	▲/—	▲/▼	—	—
Gig	▲/▲	▲/—	▲/▼	▲/▼	▲/▼
Lesson	▲/▲	▲/—	▲/—	—	—
Lifting	▲/▲	▲/—	—	—	—
Practice Songs	▲/▲	▲/—	—	—	—
Recording Session	▲/▲	▲/—	▲/▼	▲/—	▲/—
Rehearsal	▲/▲	▲/—	▲/▼	—	—
Sound and Video Check	—	—	—	—	—
Woodshedding	▲/▲	▲/—	—	—	—

Table 6.2 Effect of Success / Failure of Bass Activity on Player Stats

Chapter 7

Conclusion

7.1 Summary

This thesis presented a proposed high-level design of an electric bass tutoring system which incorporates game elements. In order to justify the design, the thesis investigated what should be considered in designing this tuitional game. Previously created computer assisted musical instrument tutoring (CAMIT) systems were thoroughly reviewed to determine which technologies and approaches could be integrated into the design. A list was given of the musical elements that would be considered during an electric bass tutoring session. Methods for extracting these musical elements were described. A multimodal method using audio from a divided pickup, MIDI messages from a bass-to-MIDI system, and computer visions was suggested as the optimal method for extracting these musical elements from an electric bass performance. A list of game elements that can be used to increase engagement was specified. Finally, how this research could be implemented in an electric bass tutoring system was outlined.

7.2 Future Work

The next natural step would be to implement the design. Implementing the design may be expensive and take quite a while. Therefore, several design requirements would need to be removed first and added on in later versions of the implementation. After the software is developed it could be tested to determine its effectiveness in increasing engagement.

Some design requirements that could also be added include the system's ability to use computer vision to evaluate the stage performance of the bassists. Stage presence is an important part of musicians' performing skills. Therefore, in order for the system to help players improve their stage performance skills it would need to be able to *see* how the students' bodies move. The use of computer vision for gesture recognition, such as the technology used by the Microsoft Xbox Kinect, would need to be investigated.

Another future work would be to extend the game design to different instruments. This would be especially useful if internet connection latency decreases to a point where it would be possible to have, for example, a guitarist and a bassist playing a song through the game *together* at the same time. This could lead to an entire virtual community sharing musical ideas and playing music together.

Bibliography

Abeßer, Jakob, Hanna Lukashevich, and Gerald Schuller (2010). "Feature-Based Extraction of Plucking and Expression Styles of the Electric Bass Guitar". In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2290–2293.

Aucouturier, Jean-Julien and François Pachet (2003). "Representing Musical Genre: A State of the Art". In: *Journal of New Music Research* 32.1, pp. 83–93.

Aucouturier, Jean-Julien and Mark Sandler (2001). "Segmentation of musical signals using hidden Markov models". In: *Proceedings of the 110th Audio Engineering Society Convention*. Amsterdam, The Netherlands.

AXON Technologies (2010). *AXON Technologies*. URL: `http://www.axon-technologies.net/` (visited on 08/04/2010).

Bartle, Richard (1996). *Hearts, Clubs, Diamonds, Spades: Players Who Suit MUDs*. URL: `http://www.mud.co.uk/richard/hcds.htm` (visited on 08/10/2010).

Bello, JP, L Daudet, S Abdullah, C Duxbury, M Davies, and M Sandler (2005). "A Tutorial on Onset Detection in Music Signals". In: *IEEE Transactions on Speech and Audio Processing* 13.5.

Bevilacqua, Frédéric, Nicolas Rasamimanana, Emmanuel Fléty, Serge Lemouton, and Florence Baschet (2006). "The Augmented Violin Project: Research, Composition and Performance Report". In: *Proceedings of 6th International Conference on New Interfaces for Musical Expression (NIME 06)*. Paris, France.

Brandão, Márcio, Geraint Wiggins, and Helen Pain (1999). "Computers in Music Education". In: *Proceedings of the AISB'99 Symposium on Musical Creativity*. AISB.

Burns, Anne-Marie (2006). "Computer Vision Methods for Guitarist Left-Hand Fingering Recognition". MA thesis. McGill University.

Burns, Anne-Marie and Barbara Mazzarino (2005). "Finger Tracking Methods Using EyesWeb". In: *6th International Gesture Workshop (GW05)*. Vannes, France, pp. 156–167.

Burns, Anne-Marie and Marcelo M. Wanderley (2006a). “Computer Vision Method for Guitarist Fingering Retrieval”. In: *Sound and Music Computing 06*. Marseille, France.

— (2006b). “Visual Methods for the Retrieval of Guitarist Fingering”. In: *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME06)*. Paris, France.

Cakmakci, Ozan, François Bèrard, and Joëlle Coutaz (2003). “An augmented reality based learning assistant for electric bass guitar”. In: *Proceedings of the 10th International Conference on Human-Computer Interaction*. Crete, Greece.

Celemony (2010). *Celemony*. URL: `http://www.celemony.com` (visited on 07/27/2010).

Chang, Wei-Chen, Alvin W. Y. Su, Chunghsin Yeh, Axel Roebel, and Xavier Rodet (2008). “Multiple-F0 tracking based on a high-order HMM model”. In: *Proceedings of the International Conference on Digital Audio Effects (DAFx-08)*, pp. 379–386.

Chen, Hung-Chen, Chih-Hsiang Lin, and Arbee L.P. Chen (2004). “Music Segmentation by Rhythmic Features and Melodic Shapes”. In: *Proceedings of the 2004 s International Conference of Multimedia and Expo (ICME)*, pp. 1643–1646.

Cheveigné, Alain de and Hideki Kawahara (2002). “YIN, a fundamental frequency estimator for speech and music”. In: *The Journal of the Acoustical Society of America* 111, pp. 1917–1930.

Cont, Arshia (2006). “Realtime Multiple Pitch Observation Using Sparse Non-Negative Constraints”. In: *Proceedings of the International Conference on Music Information Retrieval (ISMIR06)*.

— (2010). “A Coupled Duration-Focused Architecture for Real-Time Music-to-Score Alignment”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.6, pp. 974–987.

Cont, Arshia, Shlomo Dubnov, and David Wessel (2007). “Realtime Multiple-pitch and Multiple-instrument Recognition For Music Signals using Sparse Non-negative Constraints”. In: *Proceedings of Digital Audio Effects Conference (DAFx)*. Bordeaux.

Cooper, Helen (1985). Berkley Publishing Group.

Csikszentmihalyi, Mihaly (1975). *Beyond Boredom and Anxiety*. 1st ed. Jossey-Bass, San Francisco.

— (1990). *Flow : The Psychology of Optimal Experience*. English. 1st ed. Harper & Row, New York.

Cuadra, Patricio De La and Aaron Master (2001). "Efficient pitch detection techniques for interactive music". In: *Proceedings of the 2001 International Computer Music Conference, La Habana*.

Dannenberg, Roger B., Marta Sanchez, Annabelle Joseph, Robert Joseph, Ronald Saul, and Peter Capell (1990). "A Computer-Based Multi-Media Tutor for Beginning Piano Students". In: *Journal of New Music Research* 19.2-3, pp. 155–173.

— (1993). "Results from the Piano Tutor Project". In: *Proceedings of the Fourth Biennial Arts and Technology Symposium*, pp. 143–150.

Dervan, Siobhán and Bill McDaniel (2007). "Practice Perfect - An Intelligent Music Tutoring System for Assisting and Improving Practice and Exam Preparation". In: *Proceedings of Ed-Media, AACE*.

Dima, Gabriel and Alina Borcos (2008). "Remote Music Teaching Based on the Use of Advanced ICTs". In: *Proceedings of the Bridging the Digital Divide in Rural Communities Conference*.

Dixon, Simon (2001). "Automatic Extraction of Tempo and Beat from Expressive Performances". In: *Journal of New Music Research* 30, pp. 39–58.

Doidge, Norman (2007). *The Brain That Changes Itself: Stories of Personal Triumph from the Frontiers of Brain Science*. Penguin Books.

Dondlinger, Mary Jo (2007). "Educational Video Game Design: A Review of the Literature". In: *Journal of Applied Educational Technology* 4.1.

Dosenbach, Kerstin, Wolfgang Fohl, and Andreas Meisel (2008). "Identification of Individual Guitar Sounds by Support Vector Machines". In: *Proceedings of the International Conference on Digital Audio Effects (DAFx-08)*, pp. 317–320.

EBow (2011). *EBow*. URL: `http://www.ebow.com/` (visited on 08/04/2010).

Rock Band, Harmonix (2007). *Rock Band*. URL: `http://www.rockband.com` (visited on 01/20/2011).

Evangelista, Gianpaolo and Fredrik Eckerholm (2010). "Player-Instrument Interaction Models For Digital Waveguide Synthesis of Guitar: Touch and Collisions". In: *IEEE Transactions on Audio, Speech, and Language Processing* 18 (4).

EVE Online (2003). *EVE Online*, CCP Games, first released in 2003. URL: `http://www.eveonline.com/` (visited on 08/10/2010).

Fender (2011). *Squier by Fender Stratocaster Guitar and Controller*. URL: `http://www.fender.com/promos/2010/rockband3/thecontroller.php` (visited on 01/20/2011).

Fober, Dominique, Stephane Letz, and Yann Orlarey (2005). "IMUTUS Score Processing Components". In: *Proceedings of the International Computer Music Conference*. Ed. by ICMA, pp. 427–430.

— (2007). "VEMUS - Feedback and Groupware Technologies for Music Instrument Learning". In: *Proceedings of the 4th Sound and Music Computing Conference SMC'07 - Lefkada, Greece*, pp. 117–123.

Fober, Dominique, Stephane Letz, Yann Orlarey, Anders Askenfeld, Kjetil Falkenberg Hansen, and Erwin Schoonderwaldt (2004). "IMUTUS - an Interactive Music Tuition System". In: *Proceedings of the First Sound and Music Computing conference - SMC'04*. Ed. by IRCAM, pp. 97–103.

Foote, Jonathan T. (2000). "Automatic audio segmentation using a measure of audio novelty". In: *Proceedings of IEEE International Conference on Multimedia and Expo*. Vol. 1, pp. 452–455.

Friedland, Ed (1996). *Hal Leonard Bass Method - Complete Edition: Books 1, 2 and 3 Bound Together in One Easy-to-Use Volume*. Hal Leonard Corporation.

Frisson, Christian, Loïc Reboursière, Wen-Yang Chu, Otso Lähdeoja, John Anderson Mills III, Cécile Picard, Ao Shen, and Todor Todoroff (2009). "Multimodal Guitar: Performance Toolbox and Study Workbench". In: *QPSR of the numediart research program*. Ed. by Thierry Dutoit and Benoît Macq. Vol. 2. 3. numediart Research Program on Digital Art Technologies, pp. 67–84. URL: `http://www.numediart.org/docs/numediart_2009_s07_p1_report.pdf`.

Garris, Rosemary, Robert Ahlers, and James E. Driskell (2002). "Games, Motivation, and Learning: A Research and Practice Model". In: *Simulation & Gaming* 33, pp. 441–467.

Gee, James Paul (2003). *What Video Games Have to Teach Us About Learning and Literacy*. New York: Palgrave Macmillan.

Gerhard, David (2003). *Pitch Extraction and Fundamental Frequency: History and Current Techniques*. Tech. rep. University of Regina.

Gladwell, Malcolm (2005). *Blink, The Power of Thinking Without Thinking*. Back Bay Books.

Goodwin, Michael M. and Jean Laroche (2003). "Audio Segmentation By Feature-Space Clustering Using Linear Discriminant Analysis and Dynamic Programming". In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 131–134.

Goto, Masataka (2001). "An Audio-based Real-time Beat Tracking System for Music With or Without Drum-sounds". In: *Journal of New Music Research* 30 (2), pp. 159–171.

Graph Tech Guitar Labs (2010). *Graph Tech Ghost Modular Pickup Systems*. URL: `http://www.graphtech.com/products.html?CategoryID=2` (visited on 08/04/2010).

Hollinger, Diana Marie (2007). "Instrument of Social Reform: A Case Study of the Venezuelan System of Youth Orchestras". PhD thesis. Arizona State University.

Howard, David and Jamie Angus (2001). *Acoustics and Psychoacoustics*. 2nd ed. Focal Press.

Huron, David (1996). "The Melodic Arch in Western Folksongs". In: *Computing in Musicology* 10, pp. 3–23.

i-Maestro (2010). *i-Maestro*. URL: `http://www.i-maestro.org/` (visited on 06/18/2010).

IBM (2010). *Lotus product lifecycle dates*. URL: `http://www-01.ibm.com/software/lotus/support/lifecycle/` (visited on 08/01/2010).

IMUTUS (2010). *IMUTUS Resources and Newsletters*. URL: `http://www.exodus.gr/imutus/resources.htm` (visited on 06/20/2010).

Industrial Radio (2010). *Industrial Radio Midi Bass*. URL: `http://www.industrialradio.com.au/` (visited on 08/04/2010).

Iyengar, Sheena S. and Mark R. Lepper (1999). "Rethinking the Value of Choice: A Cultural Perspective on Intrinsic Motivation". In: *Journal of Personality and Social Psychology* 76.3, pp. 349–366.

Jisi, Chris (2003). *Brave New Bass : Interviews and Lessons With The Innovators, Trendsetters and Visonaries*. Backbeat Books.

Juslin, Patrik N. (2004). "Computer Feedback in Musical Communication of Emotions: The Feel-ME program". In: *Proceedings of the 18th Congress of the International Association of Empirical Aesthetics*, pp. 550–554.

Juslin, Patrik N. and Petri Laukka (2003). "Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?" In: *Psychological Bulletin* 129.5, pp. 770–814.

Juslin, Patrik N., Jessika Karlsson, Erik Lindström, Anders Friberg, and Erwin Schoonderwaldt (2006). "Play It Again With Feeling: Computer Feedback in Musical Communication of Emotions". In: *Journal of Experimental Psychology* 12.2, pp. 79–95.

Karlsson, Jessika, Erik Lindström, and Patrik N. Juslin (2009). "Teaching Musical Expression: Effects of Production and Delivery of Feedback by Teacher vs. Computer on Rated Feedback Quality". In: *Music Education Research* 11.2, pp. 175–191.

Kato, Hirokazu and Mark Billinghurst (1999). "Marker tracking and HMD calibration for a video-based augmented reality conferencing system". In: *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality*, pp. 85–94.

Kennedy, Michael and Joyve Bourne (2006). *The Oxford Dictionary of Music*. 2nd ed. revised. Oxford University Press.

Kerdvibulvech, Chutisant and Hideo Saito (2008). "Guitarist Fingertip Tracking by Integrating a Bayesian Classifier into Particle Filters". In: *International Journal of Advances in Human-Computer Interaction (AHCI)*.

Kiili, Kristian (2005). "Digital Game-Based Learning: Towards An Experiential Gaming Model". In: *Internet and Higher Education* 8 (1), pp. 13–24.

Kim, Amy Jo (2009). *Putting the Fun in Functional: Applying Game Mechanics to Social Media, a presentation for Startup2Startup Gathering, 2009*. URL: `http://www.slideshare.net/amyjokim/fun-in-functional-2009-presentation` (visited on 07/25/2010).

— (2010). *MetaGame Design: Reward Systems That Drive Engagement, a presenation for Game Developers Conference, 2010*. URL: `http://www.gdcvault.com/play/1012242/Meta-Game-Design-Reward-Systems` (visited on 07/25/2010).

Klapuri, Anssi (2006). "Auditory Model-Based Methods for Multiple Fundamental Frequency Estimation". In: *Signal Processing Methods for Music Transcription*. Ed. by Anssi Klapuri and Manuel Davy. New York: Springer, pp. 229–265.

Klapuri, Anssi P., Antti J. Eronen, and Jaakko T. Astola (2004). "Analysis of the Meter of Acoustic Musical Signals". In: *IEEE Trans. Speech and Audio Processing*, pp. 342–355.

Knösche, Thomas R., Christiane Neuhaus, Jens Hauelsen, Kai Alter, Burkhard Maess, Otto W. Witte, and Angela D. Friederici (2005). "Perception of Phrase Structure in Music". In: *Human Brain Mapping* 24, pp. 259–273.

Krumhansl, Carol L. (2000). "Rhythm and Pitch in Music Cognition". In: *Psychological Bulletin* 126.1, pp. 159–179.

Lartillot, Olivier and Petri Toiviainen (2007). "A Matlab Toolbox For Musical Feature Extraction From Audio". In: *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-07)*. Bordeaux, France.

Latham, Alison (2002). *The Oxford Companion to Music*. Oxford University Press, USA.

Lehrer, Lonah (2009). *How We Decide*. Houghton Mifflin Harcourt.

Levitin, Daniel Joseph (2006). *This Is Your Brain On Music*. New York, NY:Penguin.

Levy, Mark (2008). “Structural Segmentation of Musical Audio by Constrained Clustering”. In: *IEEE Transactions on Audio, Speech, and Language Processing* 16.2.

Li, Dongge, Ishwar K. Sethi, Nevenka Dimitrova, and Tom McGee (2001). “Classification of General Audio Data for Content-based Retrieval”. In: *Pattern Recognition Letters* 22 (5), pp. 533–544.

Lim, Sohye and Byron Reeves (2009). “Being in the Game: Effects of Avatar Choice and Point of View on Arousal Responses During Play”. In: *Media Psycholog* 12 (4), pp. 348–370.

Lu, Huanhuan, Bingjun Zhang, Ye Wang, and Wee Kheng Leow (2008). “iDVT: A Digital Violin Tutoring System based on Audio-Visual Fusion”. In: *ACM Multimedia Conference*. Vancouver, British Columbia, Canada.

Lu, Lie and Hong-Jiang Zhang (2003). “Automated Extraction of Music Snippets”. In: *Proceedings of the Eleventh ACM International Conference on Multimedia*, pp. 140–147.

McAdams, Stephen (1999). “Perspectives on the Contribution of Timbre to Musical Structure”. In: *Computer Music Journal* 23 (3), pp. 85–102.

Mcennis, Daniel, Cory Mckay, and Ichiro Fujinaga (2006). “jAudio: Additions and Improvements”. In: *In Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR*, p. 385.

McGee, Sean (2007). *Qwerty Warrior 2. Qwerty Warrior 2*, Diseased Productions. URL: `http://www.crazymonkeygames.com/QWERTY-Warriors-2.html` (visited on 08/10/2010).

McKay, Cory (2004). “Automatic Genre Classification of MIDI Recordings”. MA thesis. McGill University.

— (2010). “Automatic music classification with jMIR”. PhD thesis. McGill University.

Mckay, Cory and Ichiro Fujinaga (2006). “jSymbolic: A Feature Extractor for MIDI Files”. In: *In Int. Computer Music Conf*, pp. 302–305.

McMillen, Keith (2010). *String Port by Keith McMillen*. URL: `http://www.keithmcmillen.com/` (visited on 08/14/2010).

Miller, Donald G. (2011). *Voce Vista*. URL: `http://www.vocevista.com` (visited on 01/20/2011).

Mills, Janet (1991). “Assessing musical performance musically”. In: *Educational Studies* 17.2, pp. 173–181.

moodle.org (2010). *Moodle.org: open-source community-based tools for learning*. URL: `http://moodle.org` (visited on 06/10/2010).

Nardo, Rachel (2010). "Got Rhythm? Get RockBand[TM]". In: *General Music Today* 23.2, pp. 38–39.

Ng, Kia (2008a). "4th i-Maestro Workshop on Technology-Enhanced Music Education". In: *Proceedings of the 8th International conference on New Interfaces for Musical Expression (NIME 2008)*. Genova, Italy.

— (2008b). "Technology-Enhanced Learning for Music with i-Maestro Framework and Tools". In: *Proceedings of EVA London 2008: the International Conference of Electronic Visualisation and the Arts*. British Computer Society, 5 Southampton Street, London WC2E 7HA, UKHA, UK.

— (2009). "i-Maestro Technology-Enhanced Learning for Music with Gesture Visualisation and Sonification". In: *Proceedings of the 2nd International Conference of the Reflective Conservatoire - Building Connections*. Guildhall School of Music, Drama, and Barbican Conference Centre, London.

Ng, Kia, Tillman Weyde, and Paolo Nesi (2008). "i-Maestro: Technology-Enhanced Learning for Music". In: *Proceedings of ICMC2008: International Computer Music Conference*. Belfast.

Ng, Kia, Tillman Weyde, Oliver Larkin, Kerstin Neubarth, Thijs Koerselman, and Bee Ong (2007). "3D Augmented Mirror: A Multimodal Interface for String Instrument learning and Teaching with Gesture Support". In: *Proceedings of the 9th International Conference on Multimodal Interfaces*, pp. 339–345.

Ng, Kia, Bee Ong, Tillman Weyde, and Kerstin Neubarth (2008). "Interactive Multimedia Technology-Enhanced Learning for Music with i-Maestro". In: *Proceedings of ED-MEDIA 2008 World Conference on Education Multimedia*. Hypermedia and Telecommunications, Vienna, Austria.

Nintendo (2010). *Super Mario Galaxy 2*. Nintendo Co., Ltd.

NMAC (2011). *MENC: The National Association for Music Education: National Music Adjudication Coalition (NMAC)*. URL: `http://www.menc.org/resources/view/national-music-adjudication` (visited on 01/20/2011).

Ong, Bee, Kia Ng, Nicola Mitolo, and Paolo Nesi (2006). "i-Maestro: Interactive Multimedia Environments for Music Education". In: *Proceedings of the Second International Conference on Automated Production of Cross Media Content for Multichannel Dis-*

tribution (AXMEDIS 2006). Vol. Volume for Workshops, Tutorials, Applications and Industrial. Leeds, UK: Firenze University Press (FUP), pp. 87–91.

Orio, Nicola, Serge Lemouton, and Diemo Schwarz (2003). "Score following: state of the art and new developments". In: *Proceedings of the 2003 conference on New interfaces for musical expression*. NIME '03. Montreal, Quebec, Canada: National University of Singapore, pp. 36–41. URL: `http://portal.acm.org/citation.cfm?id=1085714.1085724`.

Pastorius, Jaco (1985). *Modern Electric Bass (Instructional Video)*. DCI Music Video Productions, Inc.

Percival, Graham (2008). "Computer-Assisted Musical Instrument Tutoring with Targeted Exercises". MA thesis. University of Victoria.

Percival, Graham, Ye Wang, and George Tzanetakis (2007). "Effective Use of Multimedia for Computer-Assisted Musical Instrument Tutoring". In: *ACM Workshop on Educational Multimedia and Multimedia Education (EMME-07)*. Augsburg, Germany, pp. 67–76.

Pfeiffer, Patrick (2010). *Bass Guitar For Dummies*. For Dummies.

Prensky, Marc (2001). *Digital Game-Based Learning*. New York: McGraw-Hill.

Puig, Vincent, Fabrice Guédy, Michel Fingerhut, Fabienne Serrière, Jean Bresson, and Olivier Zeller (2005). "Musique Lab 2: A Three Level Approach for Music Education at School". In: *Proceedings of ICMC*. Barcelona, Spain.

Rank, Erhard and Gernot Kubin (1997). "A Waveguide Model for Slapbass Synthesis". In: *Proceedings of the 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*. Vol. 1, p. 443.

Raphael, Christopher (1999). "Automatic Segmentation of Acoustic Musical Signals Using Hidden Markov Model". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21.4, pp. 360–370.

Raptis, Spyros, Aimilios Chalamandaris, Alexandros Baxevanis, Anders Askenfelt, Erwin Schoonderwaldt Kjetil, Falkenberg Hansen, Dominique Fober, Stéphane Letz, and Yann Orlarey (2005). "IMUTUS - An Effective Practicing Environment For Music Tuition". In: *Proceedings of the International Computer Music Conference (ICMC)*. Ed. by ICMA, pp. 383–386.

RedOctane (2005). *Guitar Hero*. URL: `http://hub.guitarhero.com/` (visited on 01/20/2011).

Reeves, Byron and J. Leighton Read (2009). *Total Engagement: How Games and Virtual Worlds Are Changing the Way People Work and Businesses Compete*. Boston: Harvard Business Press.

Rieber, Lloyd P. (1996). "Seriously Considering Play: Designing Interactive Learning Environments Based on the Blending of Microworlds, Simulations, and Games". In: *Educational Technology Research and Development* 44.2, pp. 43–58.

Riemann, Hugo (1900). *Vademecum der Phrasierung.* Leipzig: Max-Hess-Verlag.

Robine, Matthias, Graham Percival, and Mathieu Lagrange (2007). "Analysis of Saxophone Performance for Computer-Assisted Tutoring". In: *Proceedings of the International Computer Music Conference (ICMC07)*. Vol. 2. Copenhagen, Denmark, pp. 381–384. URL: `http://www.labri.fr/publications/is/2007/RPL07`.

Roland Corporation (2010a). *Roland GK-3B: Divided Bass Pickup*. URL: `http://www.rolandus.com/products/productdetails.php?ProductId=648` (visited on 08/04/2010).

— (2010b). *Roland VB-99: V-Bass System*. URL: `http://www.rolandus.com/products/productdetails.php?ProductId=1000` (visited on 08/04/2010).

Rollings, Andrew and Ernest Adams (2003). *Andrew Rollings and Ernest Adams on Game Design*. New Riders Games.

Salen, Katie and Eric Zimmerman (2003). *Rules of Play : Game Design Fundamentals*. The MIT Press.

Scheirer, Eric D. (1998). "Tempo and Beat Analysis of Acoustic Musical Signals". In: *Journal of the Acoustical Society of America* 103, pp. 588–601.

Schnell, Norbert, Frederic Bevilacqua, Diemo Schwarz, Nicolas Rasamimanana, and Fabrice Guedy (2006). "Technology and Paradigms to Support the Learning of Music Performance". In: *Proceedings of the Second International Conference on Automated Production of Cross Media Content for Multi-channel Distribution (AXMEDIS 2006)*, pp. 87–91.

Schoonderwaldt, E., A. Askenfelt, and K. Hansen (2005). "Design and Implementation of Automatic Evaluation of Recorder Performance in IMUTUS". In: *Proceedings of ICMC*. Barcelona, Spain. URL: `http://www.speech.kth.se/prod/publications/files/1296.pdf`.

Schoonderwaldt, Erwin, Kjetil Falkenberg Hansen, and Anders Askenfelt (2004). "IMUTUS - An interactive system for learning to play a musical instrument". In: *Proceedings of the International Conference of Interactive Computer Aided Learning (ICL)*.

Sha, Fei and Lawrence K. Saul (2005). "Real-Time Pitch Determination of One or More Voices by Nonnegative Matrix Factorization". In: *Advances in Neural Information Processing Systems 17*. MIT Press, pp. 1233–1240.

Shaffer, David Williamson, Kurt D. Squire, Richard Halverson, and James Paul Gee (2005). "Video Games and the Future of Learning". In: *Phi Delta Kappan* 87.2, pp. 104–111.

Sing & See (2011). *Sing & See*. URL: `http://www.singandsee.com` (visited on 01/20/2011).

Smart Music (2011). Smart Music *a product of Make Music*. URL: `http://www.smartmusic.com` (visited on 01/20/2011).

Smith, Graeme and Andrew Johnston (2008). "Interactive Software for Guitar Learning". In: *Proceedings of the Australasian Computer Music Conference*, pp. 69–77.

Smoliar, Stephen W., John A. Waterworth, and Peter R. Kellock (1995). "A System For Piano Education Beyond Notation Literacy". In: *3rd ACM International Conference on Multimedia*. San Francisco, California, United States: ACM Multimedia 95, pp. 457–465.

So, Aram (2008). "Guitar man: (an implementation of a rhythm game cooperative musical performance system with actual musical instruments)". In: *SIGGRAPH Asia '08: ACM SIGGRAPH ASIA 2008 educators programme*. Singapore: ACM, pp. 1–4. DOI: `http://doi.acm.org/10.1145/1507713.1507727`.

Sonuus (2010). *Sonuus B2M – Universal Bass to MIDI Converter*. URL: `http://www.sonuus.com/products_b2m.html` (visited on 08/04/2010).

Squire, Kurt (2006). "From Content to Context: Videogames as Designed Experience". In: *Educational Researcher* 35.19, pp. 19–29.

Star Play It (2011). Star Play It. URL: `http://www.smartmusic.com` (visited on 01/20/2011).

Tambouratzis, George, Kostas Perifanos, Iro Voulgari, Anders Askenfelt, Svante Granqvist, Kjetil Falkenberg Hansen, Yann Orlarey, Dominique Fober, and Stephane Letz (2008). "VEMUS: An Integrated Platform to Support Music Tuition Tasks". In: *Proceedings of the 2008 Eighth IEEE International Conference on Advanced Learning Technologies*, pp. 972–976.

Traube, Caroline (2004). "An Interdisciplinary Study of the Timbre of the Classical Guitar". PhD thesis. McGill University.

Traube, Caroline and Philippe Depalle (2003). "Extraction of the Excitation Point Location on a String Using Weighted Least-Square Estimation of a Comb Filter Delay". In: *Proceedings of the 6th International Conference on Digital Audio Effects (DAFx-03)*.

Traube, Caroline, Philippe Depalle, and Marcelo Wanderley (2003). "Indirect Acquisition of Instrumental Gesture Based on Signal, Physical and Perceptual Information". In: *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03)*. Montreal, Quebec, Canada, pp. 42–47.

Tzanetakis, George (2002). "Musical Genre Classification of Audio Signals". In: *IEEE transactions on speech and audio processing* 10.5, pp. 293–302.

Tzanetakis, George and Perry Cook (1999). "MARSYAS: a framework for audio analysis". In: *Organised Sound* 4.3, pp. 169–175. DOI: `http://dx.doi.org/10.1017/S1355771800003071`.

— (2000). "Audio Information Retrieval (AIR) Tools". In: *Proceedings of International Symposium on Music Information Retrieval (ISMIR)*.

Vasser, Anthony (2010). *Audi-Graph*. URL: `http://www.audi-graph.com` (visited on 01/20/2011).

VEMUS (2009). *vemusFinal.wmv*. URL: `http://api.vemus.org/dl/vemusFinal.wmv` (visited on 06/10/2010).

— (2010). *Welcome to VEMUS! - VEMUS - Virtual European Music School*. URL: `http://www.vemus.org/` (visited on 06/18/2010).

Wang, Ye and Bingjun Zhang (2008). "Application-Specific Music Transcription for Tutoring". In: *IEEE MultiMedia* 15.3, pp. 70–74.

Wang, Ye, Bingjun Zhang, and Olaf Schleusing (2007). "Educational Violin Transcription by Fusing Multimedia Streams". In: *Proceedings of the International Workshop on Educational Multimedia and Multimedia Education*. Augsburg, Bavaria, Germany, pp. 57–66.

Wessel, David L. (1979). "Timbre Space as a Musical Control Structure". In: *Computer Music Journal* 3.2, pp. 45–52.

WikiPedia. *ADSR Envelope*.

Willis, Gary (2002). *101 Bass Tips: Stuff All the Pros Know and Use*. Hal Leonard Corporation.

Wilson, Katherine A., Wendy L. Bedwell, Elizabeth H. Lazzara, Eduardo Salas, C. Shawn Burke, Jamie L. Estock, Kara L. Orvis, and Curtis Conkey (2009). "Relationships Between Game Attributes and Learning Outcomes: Review and Research Proposals". In: *Simulation & Gaming* 40.2, pp. 217–266.

Wooten, Victor (2008a). *The Music Lesson: A Spiritual Search for Growth Through Music*. Berkley Pub Group.

Wooten, Victor (2008b). *Victor Wooten: Groove Workshop (Instructional DVD)*. Hudson Music.

— (2010a). *Victor Wooten Bibliography. Official Website.* URL: http://www.victorwooten.com/bio.html (visited on 08/01/2010).

— (2010b). *Victor Wooten's Bass/Nature Camp. Official Website.* URL: http://www.bassnaturecamp.com/ (visited on 08/01/2010).

Yamaha Corporation (2010a). *Yamaha B1D Divided Pickup Unit*. URL: http://www.yamaha.com/yamahavgn/CDA/ContentDetail/ModelSeriesDetail.html?CNTID=1282 (visited on 08/04/2010).

— (2010b). *Yamaha G50 Guitar MIDI Converter*. URL: http://www.yamaha.com/yamahavgn/CDA/ContentDetail/ModelSeriesDetail.html?CNTID=1143 (visited on 08/04/2010).

Yanase, Takashi and Atsuhiro Takasu (1999). "Phrase Based Feature Extraction for Musical Information Retrieval". In: *Proceedings of the IEEE Pacific Rim Conference on Communications, Computers, and Signal Processing*, pp. 396–399.

Yin, Jun, Ye Wang, and David Hsu (2005). "Digital Violin Tutor: An Integrated System for Beginner Violin Learners". In: *Proceedings of the 13th Annual ACM international conference on Multimedia*, pp. 976–985.

Yoo, Lilit and Ichiro Fujinaga (1999). "Comparative Latency Study of Hardware and Software Pitch-Trackers". In: *Proceedings of the International Computer Music Conference (ICMC99)*, pp. 36–39.

Yourself! Fitness (2004). *Yourself! Fitness*, ResponDesign. URL: http://respondesign.com/.

Zdzinski, Stephen F. and Gail V. Barnes (2002). "Development and Validation of a String Performance Rating Scale". English. In: *Journal of Research in Music Education* 50.3, pp. 245–255. URL: http://www.jstor.org/stable/3345801.