Interactive Music Science Collaborative Activities

Team Teaching for STEAM Education

Deliverable 5.4
Initial Gesture and VR tools for music interaction and co-creation

Date:	05/10/2017
Author(s):	Aggelos Gkiokas (ATHENA), Maximos Kalliakatsos-Papakostas (ATHENA), Kosmas Kritsis (ATHENA), Nancy Zlatintsi (ATHENA), Antigoni Tsiami (ATHENA)
Contributor(s):	Carlos Acosta (LEOPOLY), Robert Piechaud (IRCAM)
Quality Assuror(s):	Carlos Acosta (LEOPOLY), Renaat Frans (UCLL)
Dissemination level:	PU
Work package	WP5 – Workbench of iMuSciCA prototypes
Version:	1.0
Keywords:	gesture interaction, performance with virtual musical instruments
Description:	This deliverable reports on the initial implementation of the gesture and VR tools for music interaction and co-creation.

Executive Summary

This deliverable reports on the initial implementation of the Gesture and Virtual Reality (G-VR) tools for music interaction and co-creation. The demonstrated G-VR tools allow students to load a previously designed virtual instrument into the interaction 3D world, select the sensor for interaction (Leap Motion or Kinect) and the mode of interaction (free or using gestures) along with several music and visualisation tools (e.g. metronome, tempogram, spectrogram and tuner among others). The students can record their performance either in the form of gestures (storing information about their movement during recording) and/or in the form of audio. After recording the audio of performances, students can open these recordings in the Performance Sound Sampler (PSS) music co-creation tool, cut sample slices on the waveform of the performance, change the pitch of the selected samples, their activation times and collaborate with other students for making new music.

Links to demonstrators:

Gesture Interactions:

http://tesla.ilsp.gr:1994/3dmii.1/

Performance Sound Sampler:

http://tesla.ilsp.gr:1994/sampler.1/

Version Log			
Date	Version No.	Author	Change
04-09-2017	0.1	Aggelos Gkiokas (ATHENA)	Initial content
27-09-2017	0.2	Aggelos Gkiokas (ATHENA), Maximos Kalliakatsos-Papakostas (ATHENA), Kosmas Kritsis (ATHENA), Nancy Zlatintsi (ATHENA), Antigoni Tsiami (ATHENA)	Text input in different section
29-09-2017	0.3	Vassilis Katsouros (ATHENA), Aggelos Gkiokas (ATHENA)	Overall supervision and preparation of version to be sent for internal review
04-10-2017	0.4	Aggelos Gkiokas (ATHENA	Reviewers comments incorporated and final version produced
05-10-2017	1.0	Vassilis Katsouros (ATHENA)	Submission to the EU

Disclaimer

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Description
G-VR	Gesture and Virtual Reality
MVC	Model-View-Controller
PSS	Performance Sample Sequencer
SPA	Single-Page Application
UI	User Interface
UX	User Experience
ATHENA	ATHENA RESEARCH AND INNOVATION CENTER IN INFORMATION COMMUNICATION & KNOWLEDGE TECHNOLOGIES
UCLL	UC LIMBURG
EA	ELLINOGERMANIKI AGOGI SCHOLI PANAGEA SAVVA AE
IRCAM	INSTITUT DE RECHERCHE ET DE COORDINATION ACOUSTIQUE MUSIQUE
LEOPOLY	3D FOR ALL SZAMITASTECHNIKAI FEJLESZTO KFT
CABRI	Cabrilog SAS
WIRIS	MATHS FOR MORE SL
UNIFRI	UNIVERSITE DE FRIBOURG

1. Introduction

This deliverable describes the first version of the Gesture and Virtual Reality (G-VR) Tools for Music Interaction and Co-creation. In the context of this deliverable two tools have been developed that allow G-VR interaction with a virtual instrument and music co-creation respectively. After the student constructs the virtual instrument (as described and demonstrated in D5.1), she/he is able to import the instrument in the G-VR environment, interact with the instrument either with free interaction or with gesture-based interaction, record the performance and import the recording in the Performance Sample Sequencer environment generating new music.

The first version of Gesture and Virtual Reality (G-VR) Tools for Music Interaction and Co-creation environment is the outcome of the integration of different core-enabling technologies including Gesture and VR Multimodal Interaction ([Deliverable 4.1, First Version of Gesture and VR multimodal interaction interface](#)), Computational Models for Sound and Music Generation ([Deliverable 4.4, First Version of Computational models for sound and music generation for virtual instruments](#)), the 3D environment of Instrument Design ([Deliverable 4.3, First Version of 3D design environment for music instruments](#)) along with audio visualizations ([Deliverable 4.5, First Version of Music Visualization and Analysis Tools](#)).

2. Installation and technical requirements

The Gesture and Virtual Reality environment and the Performance Sample Sequencer are running on the Firefox Nightly web browser. Both environments should be used in a computer with speakers for listening the generated audio.

Specifically for the G-VR interaction, the Kinect and/or Leap Motion sensors are also necessary for tracking the movement of the user in order to allow free and gesture-based interaction with the virtual instrument. Software requirements and installation instructions can be found in the Sections 2 and 3 of [Deliverable 4.1](#) (First version of gesture and VR multimodal interaction interface).

3. Description of demonstrator and user manual

In the following subsections the Gesture and Virtual Reality interaction environment and the Performance Sample Sequencer are described.

3.1. Gesture and VR tools for music interaction

3.1.1. Introduction

The Gesture and Virtual Reality (G-VR) tools for music interaction allows students to use their hands (Leap-based interaction module) and arms/body (Kinect-based interaction module) in order to interact with a virtual instrument. The environment that hosts these tools provides the user with the following functionalities:

a) Selection of the desired interaction sensor (i.e. Leap Motion or Kinect).
b) Selection of the desired gestural interaction type (i.e. free or gesture-based).
c) Loading of customized virtual instruments from the which were previously designed and constructed by the user within the *Instrument Design* environment.
d) Recording and storing user's performance sessions.
e) Generate pitch reference tones, as well as detecting and displaying in real-time the pitch of produced sound tones from the performance of virtual instruments.
f) Study audio-related visualisations that derive from either real-time performances or from previously recorded sessions.

It should be noted that free interaction and gesture-based interaction are quite different performance approaches. During free interaction, the students are "embedded" within a virtual world where they can see avatars of their hands and bodies, which, on their side, interact directly with the virtual instruments, such as virtually plucking the strings or hitting the drum membrane. On the other hand, within gesture-based interaction scenarios, the students use gestures as "motion templates" for triggering musical events from the virtual instrument.

The G-VR interaction demonstrator is hosted on[1] and incorporates a central UI which from now on will be referred to as *3D Music Interaction Interface*.

3.1.2. The 3D Music Interaction Interface

The G-VR demonstrator is operated as a responsive Single-Page Application (SPA), that follows the Model-View-Controller (MVC) software architecture. Consequently, the *3D Music Interaction Interface* works as the View component within the app architecture. It is a handy UI that provides the user with efficient and rich UX. Thus, through this interface, the user is able to easily gain access to all the functionalities that are supported by the G-VR environment.

As illustrated in Fig. 3.1.1, the *3D Music Interaction Interface* is organized into five main areas. On the top left side of the window, the sensor and interaction information is shown, i.e. which sensor and interaction type is selected. On the bottom left area, the user can access the Basic Music Tools (i.e. PitchA, Metronome etc,. see Deliverable 4.5 (First Version of Music Visualization and Analysis Tools)) and Advanced Visualizations (i.e. tempogram etc., see Deliverable 4.5 (First Version of Music Visualization and Analysis Tools)). Next, the *3D Virtual World* is shown on the central area of the demonstrator window. On the top right side of the window the user can find the *Options* button (i.e.

---

[1]http://tesla.ilsp.gr:1994/3dmii.1/

≡) that opens the *Options Menu* sidebar (Fig. 3.1.2). There, the user can choose the interaction sensor and type for his performance session, hide and show the *Tools* and *Visualization* Menus, as well as a button for loading virtual instruments designed by the user. In this version of the demonstrator, only the stringed instrument is loaded by default when the *Load* button is pressed. Finally, under the Options button, lays the the *3D Music Interaction Performance Recorder*, that allows the user to record his musical performance sessions. In the current version of this demonstrator, the recorder is able to store only the last recorded session, while each new recording overwrites the previous one. Future versions will allow saving/loading recorded performances on the local drive or the iMuSciCA cloud.
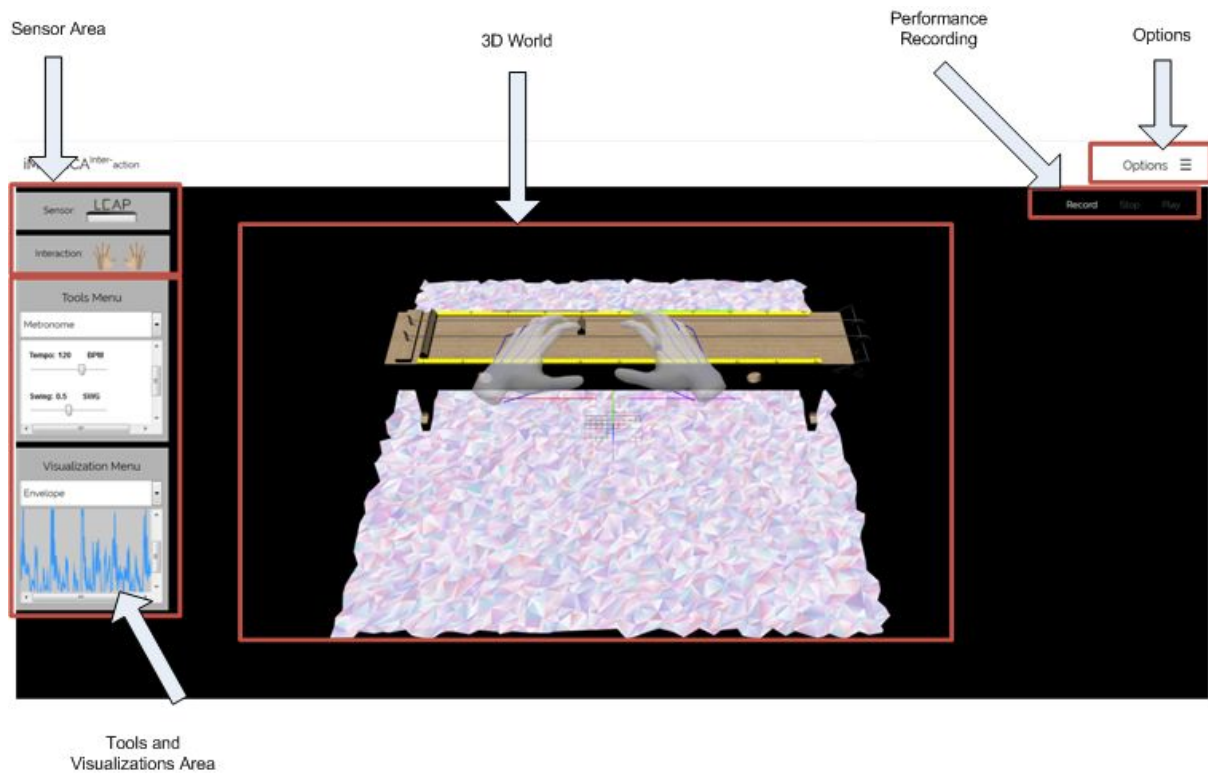

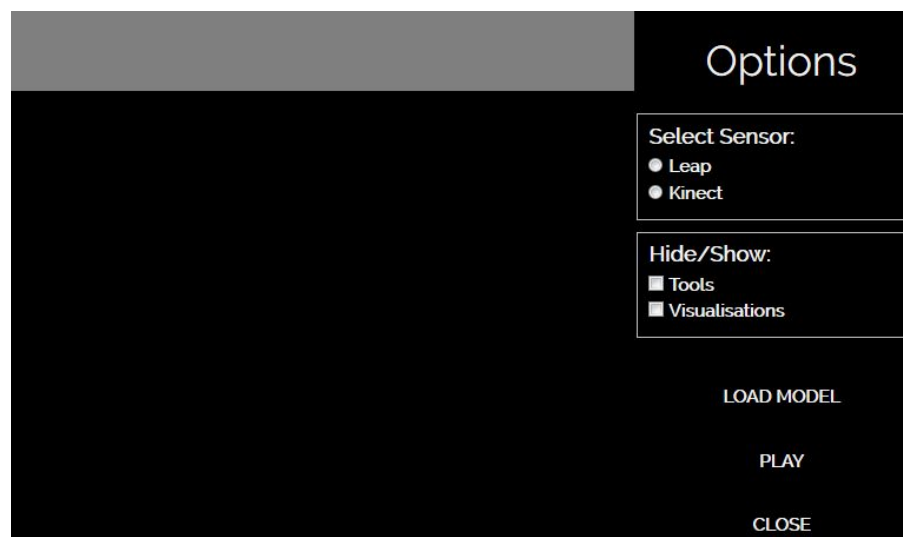**Figure 3.1.1:** Window Organization of the 3D Music Interaction Interface

Figure 3.1.2: The Options Menu

3.1.3. Selecting Sensor and Interaction Type

For selecting the interaction sensor and type, the student first needs to click on the Options button (i.e. ☰) in order to access the Options Menu. Once clicked, the *Select Sensor* field will show, which contains the two possible sensor options: the Leap Motion and the Kinect. When the user presses the radio button that corresponds to the desired interaction sensor, a new subfield should appear for selecting one of the two possible Interaction Types. As described in the Introduction (Section 3.1.1) the basic classes of interaction are *Free* and *Gesture-based Interactions* (see Fig. 3.1.3). As a visual feedback to the selected options, two fields with the respective icons are displayed at the top left area of the demonstrator window.
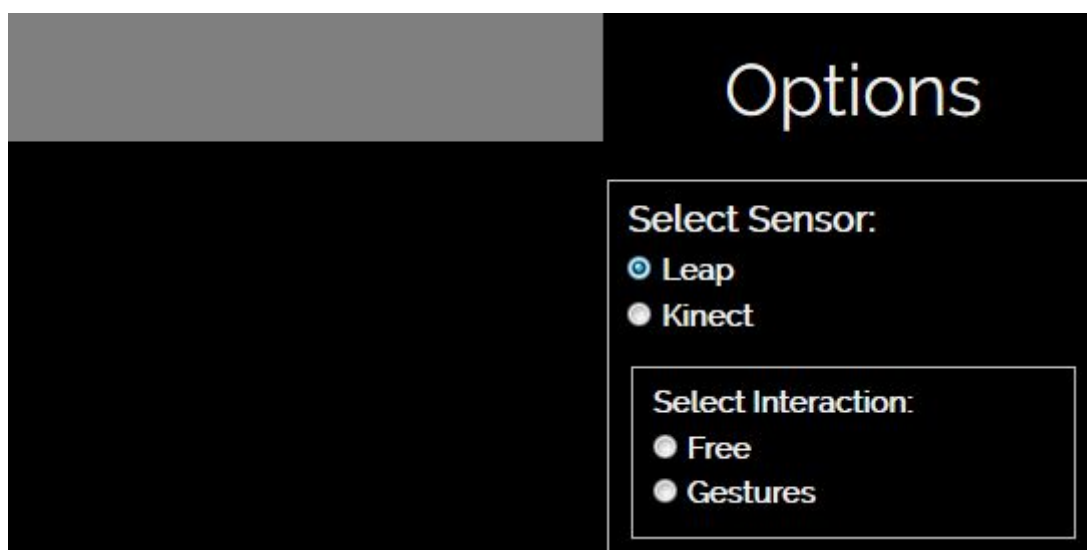

Figure 3.1.3: Sensor and Interaction Type selection menus

⚠ When the user selects Interaction Sensor and Type this cannot be undone in the current session. To use another sensor and interaction type the user must refresh the webpage discarding the cache memory of the browser.

3.1.4. Loading an Instrument

Through the 3D Music Interaction Interface the students are able to load their coveted virtual instrument with which they would interact and produce musical events. To do so, the user should firstly press the Options button (i.e. ☰) in order to open the Options menu and spot the Loading button (Fig. 3.1.4). By clicking it ,the loading process starts and once completed successfully, the 3D model of the virtual instrument should appear in the 3D virtual world. However, it is important to mention that in the current demonstrator only a string instrument is supported.
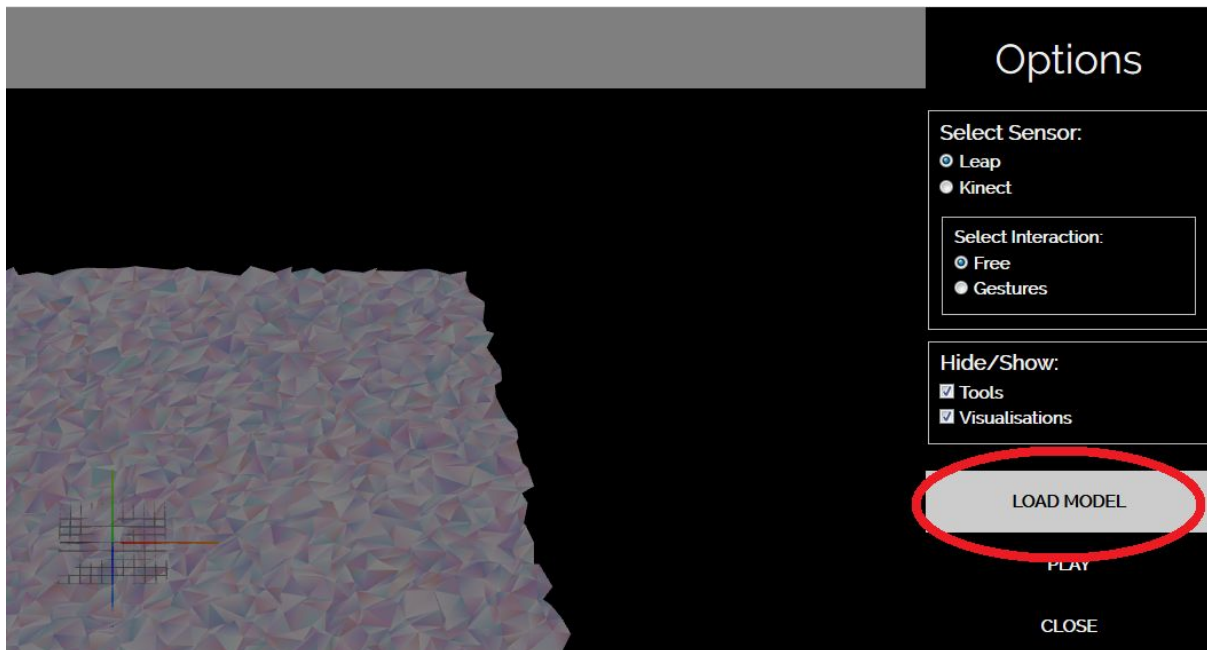
Figure 3.1.4: Load Instrument button

3.1.5. Selecting Music Tools

Figure 3.1.5: Selecting the Tools Menu

For accessing the the Music Tools area, the user first needs to press the Options button (i.e. ≡) and ensure that the respective Hide/Show field of the checkbox is checked (Fig. 3.1.5). When the user selects the Show option, the Music Tools area should appear on the left side, under the iconographic feedback of the selected Sensor and Interaction Type (see Fig 3.1.6). There, the user can select through a clickable dropdown menu several assistive tools, including a metronome, a pitch reference generator, a tuner, a musical performance audio recorder and an audio player. Once selected, the corresponding tool will appear in the Tools Menu area (Fig. 3.1.7). More information regarding these tools are given in Deliverable 4.5 (First Version of Music Visualization and Analysis Tools).
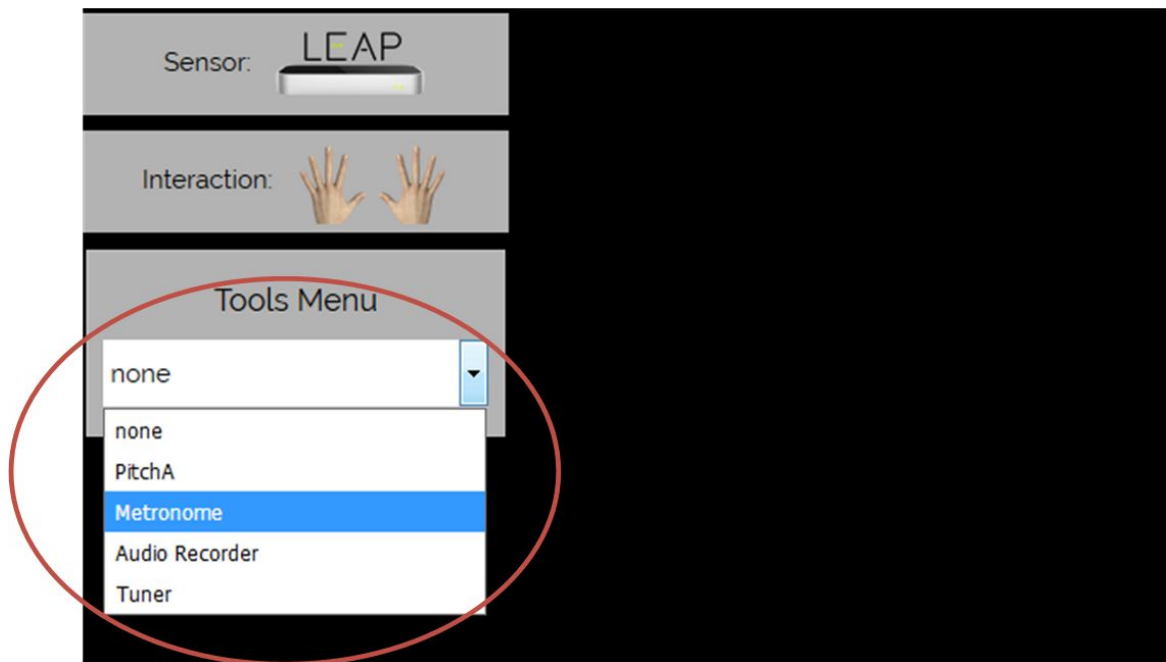
Figure 3.1.6: The Tools Menu

Figure 3.1.7: Tools menu: The PitchA tool is selected

If the user has previously made any selection in the Options, Tools and Visualization Menus, they are stored in the browser's cache memory and require to reload the website by discarding the cache memory.

3.1.6. Selecting Visualizations

Similar to the Music Tools, the user can access the Advanced Visualization Tools by clicking on the options menu and by selecting the Visualizations checkbox (see Fig. 3.1.4). Once the Visualization checkbox is selected, the list of available visualization tools will appear on the left side of the screen (Fig. 3.1.8). In the current version of the environment there are three advanced visualization types:

- Envelope
- Tempogram
- 3D Spectrogram

Details of the advanced visualization can be found on the Deliverable 4.5 (First Version of Music Visualization and Analysis Tools), Section 6.1.
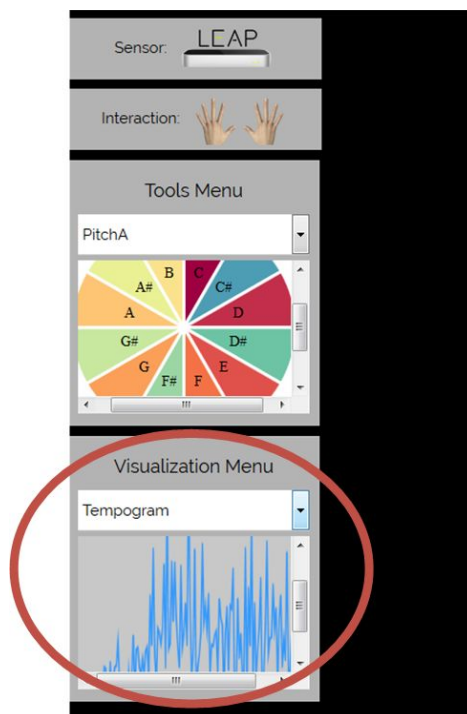

Figure 3.1.8: Tempogram visualization is selected from the Visualization Menu

3.1.7. Performing

3.1.7.1. Free Interaction Performance

In the free Interaction Performance the user is able to directly interact with the virtual instrument as being in the real world. In the case of the Leap Motion sensor, the user can see her hands in the 3D world along with the 3D instrument (Figure 3.1.9). If the fingers of the user touch the strings, then a sound corresponding to the specific string is produced by the environment based on the computational models of the instrument (Deliverable 4.4, First Version of Computational Models for Sound and Music Generation for Virtual Instruments). These models take into account the string

---

properties (e.g. length, tension, material) as well as the interaction properties (e.g. finger speed, the exact point that the string is hit etc) in order to produce the corresponding sound.
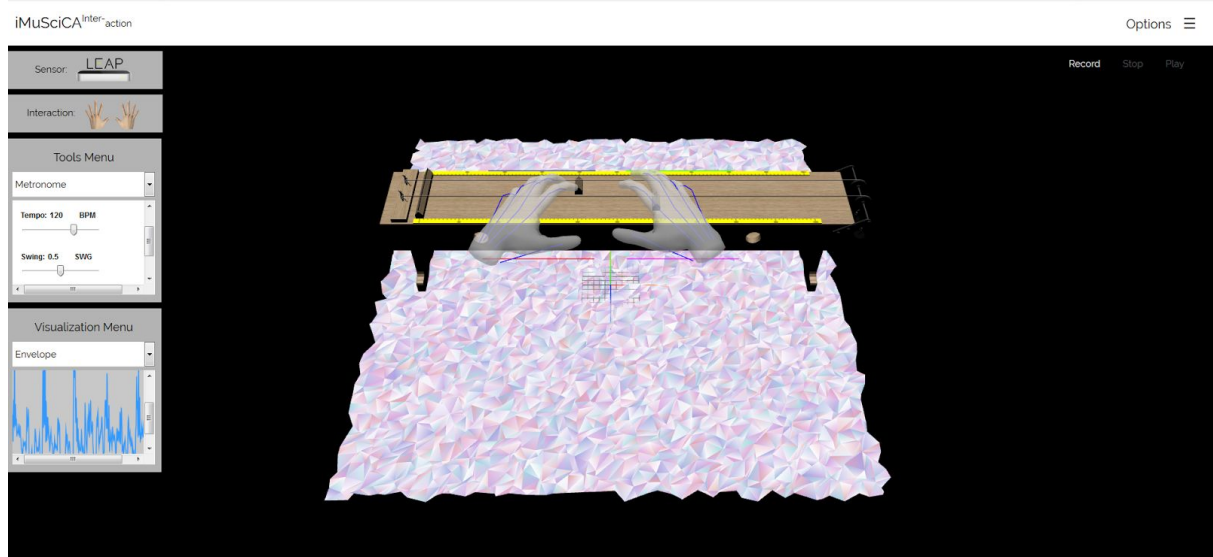

Figure 3.1.9: A screenshot from a performance session using the Leap Motion sensor and free interactions with a string instrument

In a similar manner when using the Kinect sensor the user can see his arms (Fig. 3.1.10) and can naturally interact with the virtual instrument. Note that in both cases, the user hands/arms are rescaled so that the size of the instrument and the size of arms/hands are somewhat balanced making the interaction as easy as possible.
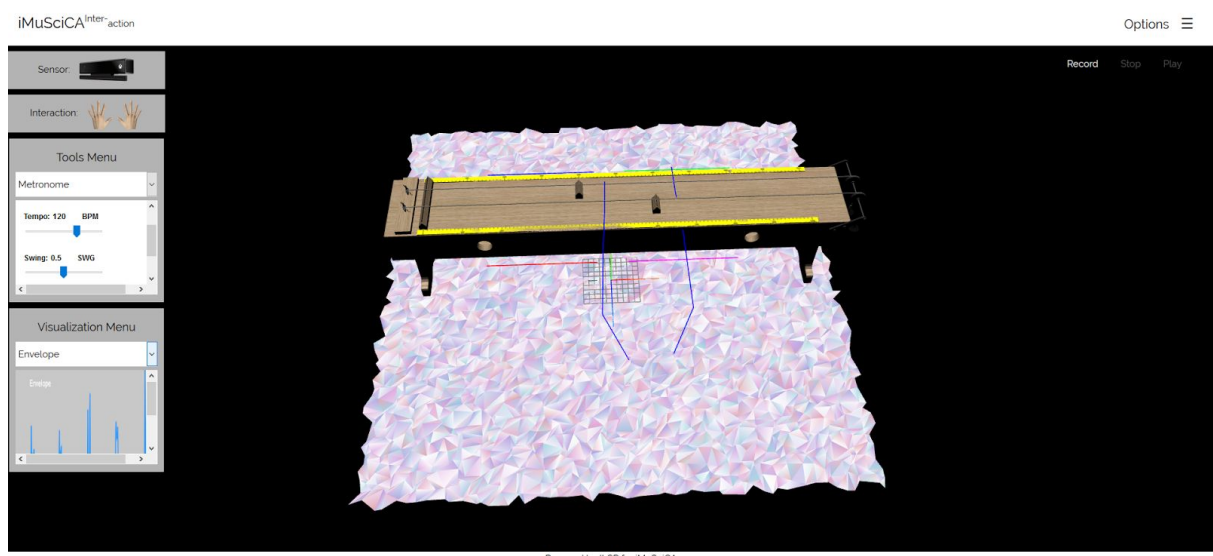

Figure 3.1.10: A screenshot from a performance session using the Kinect sensor and free interactions with a string instrument

3.1.7.2. Gesture Based Interaction Performance

Apart from the free interaction, the user can also select the Gesture-based Interaction mode for triggering sound events from the virtual instruments, by performing specific "motion templates" that have some referential similarity to the gestures that a musician follows when performing a physical instrument such as the guitar or the cello. The gesture templates depend on the selected instrument type and sensor. A detailed list of the gestures can be found on Deliverable 4.1 (First Version of Gesture and VR multimodal interaction interface). In the 3D Music Instrument Environment Interaction, we choose two type of gesture interaction, one for each of the Leap Motion/Kinect sensors.

    a. Air Guitar gesture: using the Kinect sensor, the student can perform gestures similar to the ones that a guitar-player does. This gesture is active when the student brings her/his right hand around the waist height and the left hand close to the chest height - like getting ready to perform on a guitar. When the gesture is activated, the right hand moves vertically on the waist height, simulating the pluck movement while the left hand moves vertically from the head to the waist height, simulating the movement of the hand on a virtual fretboard.

    b. Finger tapping gesture: using the Leap Motion sensor, the student can select to perform tapping-like gestures that are activated when the user puts her/his hand within the capturing range of the Leap Motion sensor. When the tapping gesture is activated, each finger of the student's right hand is assigned to a string. By doing tapping gestures, the user is able to trigger sound-generation to each respective string, while the left hand controls the pitch of that string, simulating thus a guitar fretboard.

After selecting the desired Sensor and Interaction type in the Options menu, the appropriate visualizations and music tools, and after loading the 3D Virtual World along with the virtual instrument that he would like to perform, the user can finally press the Play button in the Options area and the musical performance session should initiate successfully and start. An example screenshot is shown in Figure 3.1.11.
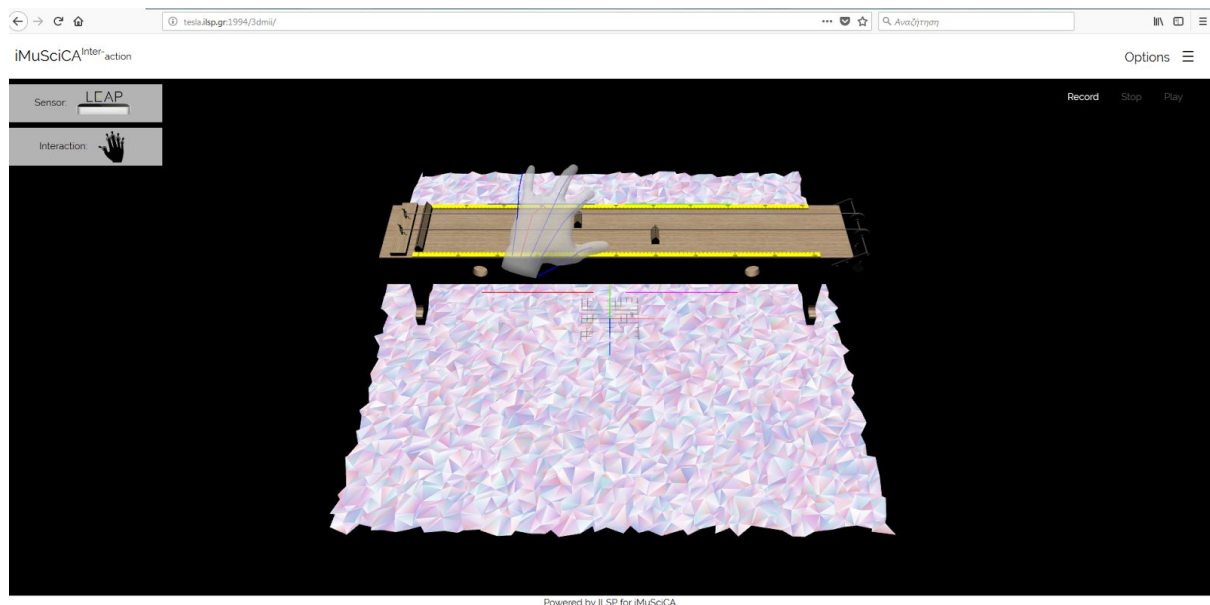
Figure 3.1.11: A screenshot from a performance session using the Leap Motions sensor and gesture interactions with a string instrument

3.2. Music co-creation: the Performance Sample Sequencer

The Performance Sample Sequencer (PSS) is a tool that allows students to work on a performance recording, obtained from the Performance/Interaction with Musical Instrument Using Gestures activity environment (see previous section 3.1)*. Students can use this recording to generate new music. To this end, stored recorded performances can be imported in audio form into the Sampler's interface. The student can then select samples from this performance recording as regions on the waveform and activate those samples on the sampler's matrix. The user can store the current settings of her/his composition and share it with other students, generating the conditions for co-creation. The PSS module is hosted online on the ATHENA server[2].

The students can interact with the interface, illustrated in Figure 3.2.1, and set the following aspects of the composition (according to the numbering presented in the aforementioned figure):

1. Select an audio file generated during performance with the instrument or with other means (in the current version, audio files stored on the server can be used).
2. Click-drag on the waveform to select up to 12 sample areas on the waveform.
3. Listen to the recording.
4. Change the tempo and the swing feel of the playback and start/stop the sampler.
5. Select time signature and the beat resolution.
6. See the automatically detected pitch (accompanied by the algorithm's confidence for the detection) of the currently selected sample and pitch-shift it between -12 and 12 semitones.
7. Either generate a random sample matrix (see 8) with random selections and activations for a random number of samples with the student-defined density and syncopation values, or select density and syncopation for generating events for separate samples (see 9).

---

[2] http://tesla.ilsp.gr:1994/sampler.1/

8. Select the activations of each sample in time by activating the respective rows and columns of the sampler matrix.
9. Generate random "Events" (according to the density and syncopation selected in 7) or sample areas (shown in 2) for each sampler separately.
10. Remove or add samples (up to 12).

Figure 3.2.1: First version of the Performance Sample Sequencer co-creation tool interface