SONIFICATION OF BOWING FEATURES FOR STRING INSTRUMENT TRAINING

Oliver Larkin, Thijs Koerselman, Bee Ong, and Kia Ng

ICSRiM – University of Leeds,
School of Music and School of Computing, Leeds LS2 9JT, UK
info@i-maestro.org www.i-maestro.org

ABSTRACT

This paper presents work on an auditory display for use in string instrument training, based on 3D motion analysis. We describe several sonifications that are intended to provide both real-time and non real-time feedback about bowing technique. Tests were conducted with string players to assess the effectiveness of the sonifications. We discuss our findings as well as ideas for further work in this area.

1. INTRODUCTION

There is a large body of research in the field of computer music devoted to augmenting string instruments and studying string performance with modern technologies such as sensors and motion capture devices [1]. One of the goals of the i-Maestro EC IST project (www.i-maestro.org) is to use these new technologies to develop innovative tools for string instrument training [2][3][4]. This is an area of music teaching that is typically untouched by technology; however, we believe that new tools based on gesture analysis and multimodal feedback could be extremely useful and beneficial in certain situations. Bowed string instruments are notoriously challenging to learn due to the precise physical control required to produce a good sound. New technologies can provide information to student and teacher on aspects of the performance that are otherwise difficult or impossible to access. For example the study in [5], using an augmented violin and motion capture system, shows how characteristics of bow-stroke gesture were found to change depending on the frequency of the bowing motion.

This paper discusses a sonification module for the i-Maestro "3D Augmented Mirror" (AMIR) [2][4], which is an application developed by the authors to facilitate the study of string performance using motion capture. AMIR is based on the traditional function of the mirror in string instrument training; the student uses it to obtain information about their performance that is not available to them from the perspective of their typical playing posture. Using 3D motion capture it is possible to enhance this functionality by viewing the performance from multiple perspectives. In addition, features can be extracted from the motion data that are not necessarily perceivable using the human eye, such as changes in bow velocity and acceleration. The AMIR application is built around a real-time interface to a motion capture system and allows synchronised recording of audio, video and motion data. The performance is visualised in 3D, accompanied by a number of graphical displays that show various analyses of features of the performance (see Figure 1).

Graphical displays are one way of delivering information about the performance; however, auditory displays offer a number of potential advantages in the context of an instrumental teaching/study aid. Auditory display can be used in "eyes-busy" situations, such as when reading a musical score [6][7]. Also, musicians are already familiar with the sound parameters typically used to convey information in auditory displays [6]. Sound may be particularly appropriate for displaying bowing gesture information since auditory perception offers better temporal resolution than visual perception. Sonification may be more appropriate than visualisation where multiple streams of information need to be communicated to the user simultaneously. Research in this field has implications for the development of enabling technologies for blind string players.

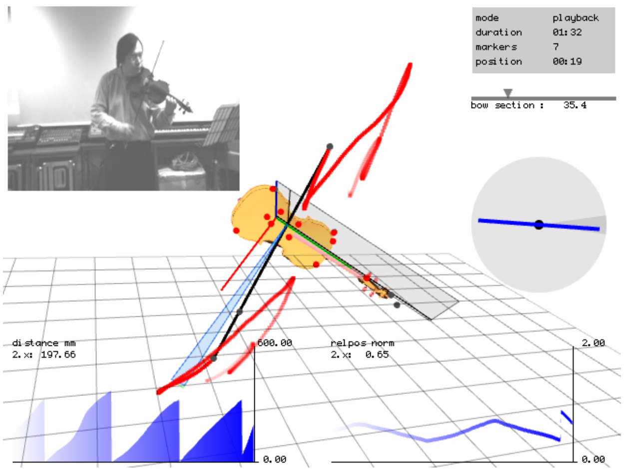

Figure 1. *Screenshot of the AMIR user interface showing graphs of bowing features.*

The application of auditory display in music instrument training is an under-explored area. One challenge is how to use sound parameters to convey information without conflicting with the sound of the instrument. When an instrumentalist is performing, their auditory perception is extremely important. This is particularly so in the case of performers who have to gauge their intonation (such as singers and string players), and who may do so using the pitch of the sound they are producing.

1.1. Hypothesis

On commencing this work our hypothesis was that for real-time sonification during musical performance, all but the simplest of sonifications would be too distracting. More complex auditory display may be useable in a non-real-time context.

2. SYSTEM

AMIR has been developed using Cycling 74's Max MSP / Jitter (www.cycling74.com) which we have extended with a library of external objects called the Motion Analysis and Visualisation (MAV) framework [4]. The sonification module is using

IRCAM's FTM library [8] as well as the standard audio synthesis and processing functionality of MSP.

Currently we are working with a VICON 8i Optical Motion Capture system (www.vicon.com), with 12 cameras and a resolution of 200 fps.

3. EXTRACTING BOWING FEATURES

AMIR provides analyses of a number of bowing features. Some of the features are appropriate for sonification, whilst some are more suited for visualisation. Some are appropriate for real-time feedback, whilst others have more potential when "reviewing" the performance after it has taken place. For bowing technique we perform analyses based on the interactions between the bow and the instrument. In order to do this it is necessary to derive a local coordinates system, which we place on the bridge of the instrument (see [2]). Once we have this transformation, the player can move freely in the motion capture stage, without affecting our analyses.

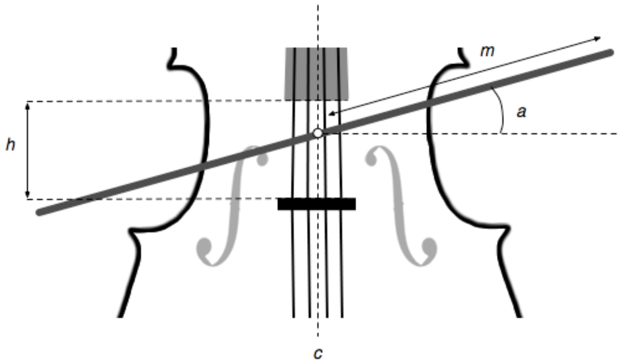

Figure 2. Bowing Analyses

Features we currently extract for sonification are:

- The *bowing segmentation*. We detect changes of bowing direction and when the bow is lifted off the strings. This segmentation allows us to look at other features on a per stroke basis. Once a recording of the performance has been made, it is possible to navigate through the recording bow stroke by bow stroke, or to look at features for strokes that match a certain criteria (e.g. all up-strokes).
- The *bow-bridge angle* (Figure 2 - *a*). One challenge of playing a string instrument is to bow parallel to the bridge, which generally produces the best sound with the least effort.
- The *bowing height* (*h*). This is the point at which the bow crosses the centre plane of the instrument (*c*) on the "fingerboard axis", normalised between the position of the bridge and the base of the fingerboard. (See Figure 2). Bowing in different positions produces different tonal qualities (e.g. "tasto" – near fingerboard, "ponticello" – near bridge).
- The *bowing magnitude (m)* and *bow position*. The *bowing magnitude* is the name that we give to the distance in millimetres from the bow tip to the contact point. The *bow position* is the *bowing magnitude* normalised to the playable length of the bow (the full length of the bow hair).
- Bow velocity and acceleration. These are calculated by taking the first and second derivatives of the *bowing magnitude*. In contrast to taking these measurements from the raw marker data this gives us values that are independent of the angle of the bow.
- *Bowing displacement curve*. This feature indicates the difference in distance travelled over the bow stroke compared to a bow stroke with a constant velocity.

4. APPROACHES TO SONIFICATION

In our initial experiments with sonification of bow gesture analysis data, we attempted to sonify continuously changing parameters in real-time, synchronous to the performance. We linked the angle of the bow to different parameters of a synthesis patch and also investigated processing the sound of the instrument by applying distortions and harmonisation to reflect changes in the extracted features. In line with our hypothesis, we quickly found that when performing, continuous real-time auditory feedback was too distracting to the performer and difficult for them to interpret simultaneously with the instrument sound. Players reported that their attention would be drawn to the sonification rather than the sound of their instrument. For these reasons we decided not to pursue this kind of auditory feedback. We found that a much more useful and appropriate modality for real-time feedback is to inform the player using the simplest of auditory displays: an auditory alert/notification. We believe it is possible that with practice a performer could adapt to play with more complex real-time sonifications; however in our opinion this would deviate too far from the traditional mode of interaction in string instrument training, and hence would not be favoured by many teachers. Another possibility is that real-time sonification can be used if specific musical exercises are designed which take into account the conflict between instrument sound and sonification, such as those described in [6].

4.1. Auditory Alerts

We allow the user to set thresholds to define acceptable ranges for analysis parameters and when these thresholds are crossed an alert is sounded. This informs the performer that certain criteria have been met with minimum distraction. The character of the notification can be changed depending on how frequently the threshold has been crossed. For example, if a player continuously repeats an error, the sound will increase in volume each time. If they start to improve and correct their error, the volume of the alert will decrease. We have used alert sounds to monitor the *bow-bridge angle*, the *bowing height* and the *bow position*. In each case the user defines the acceptable range for the parameter using a visual interface (see Figure 3). A selection of suitable audio samples is provided for the alert sounds, and different sounds can be used for different parameters.

Another auditory notification we use indicates a change of bowing direction with a percussive "click" sound (the sounds used for up and down strokes are slightly different, and are panned left and right respectively). In a situation were the student or teacher is reviewing a recording of a fast-tempo piece, they may slow down the playback and listen to the timing of the bow stroke gestures at a more appropriate speed.

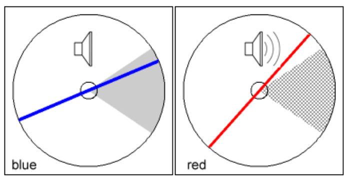

Figure 3. Visual interface for setting the *bow-bridge angle* threshold. The acceptable range is indicated by the grey area.

4.2. Sonified Line Graphs

For non real-time analysis of the performance we are exploring the use of sonified line graphs [9] to display information about analysis features for each bow stroke. Sonified line graphs have been shown to be useful in a number of different situations and can help the visually impaired to interpret trends in data sets that would otherwise be difficult to access [9][10].

To create the line graph, the continuous analysis parameters listed above are split into variable length 1D arrays based on the *bowing segmentation*, with each element in the array representing the analysis for one frame of motion data. For instance, a bow stroke that lasts one second will be represented by an array of two hundred elements when captured at 200fps. The line graph is visualised, on top of a waveform representation of the part of the audio recording that corresponds to that bow stroke. The graph may be sonified synchronous to playback or, alternatively, the user can step through the piece stroke by stroke.

We have chosen to use pitch modulation of a synthesised waveform to sonify the line graph. For each bow stroke played in the performance, a fixed pitch tone is sounded from the left channel and a modulated pitch tone is sounded from the right. Spatial separation such as this should help the user to distinguish the two tones [11]. In this way it is possible to hear the degree of deviation from the base pitch, which can help the user to perceive the deviation of the analysis feature. We use a table-lookup oscillator containing the waveform described in [12] which has been found to be particularly suitable for distinguishing pitch differences and easier to listen to than pure sine tones. An additive synthesis waveform was chosen over MIDI instrument sounds (as recommended in [13]) for a number of reasons. Synthesising the tone in our own software guarantees a stable fundamental frequency is also more precisely controllable in terms of note onset and duration, which is important for sonification that is synchronised with the playback of a recording.

The length of the tone can be adjusted so that it either equals the duration of the bow stroke, or is a fixed duration for each stroke. The first option means that the feature in question can be matched temporally to the music being played, where as the second option will give a more uniform representation of the stroke in comparison to other strokes. It should be noted that the playback of the recording may be slowed down and the duration of the sonified line graph is adjusted accordingly. This makes it possible to assess very fast strokes.

Whilst sonification of the line graph is useful to show how the analysis varies during the bow stroke, in some cases this could provide too much information. By taking the standard deviation of the values from the bow stroke array it is possible to give a more general assessment of the feature for that stroke. For instance, in the case of the bowing angle, the standard deviation can provide an indication of the overall degree of difference from a parallel bow. In this case the sonification is simplified since the second tone is not modulated over the course of the bow stroke.

The user may set the base pitch of the tone, for example they might set it to the tonic of the key of the passage that they are playing. The mapping of the modulation can be changed, depending on the resolution at which the user wishes to study the performance; for example a very determined player could set the mapping to an extreme level in order to help him/her develop precise control of bowing movements.

To calculate the *bowing displacement curve*, we first accumulate the velocity of each frame in the bow stroke. This gives us (*g(t)*) (2) which represents the distance travelled over the stroke. Next we subtract a normalised linear function (*f(t)*)(1), which leaves us (*s(t)*) representing the displacement curve in relation to a linear-velocity bow stroke (3) (see Figure 4).

$$f(t) = t, t \geq 0 \quad (1)$$

$$g(t) = \left| d_{t+1} - d_t \right|, t \geq 0 \quad (2)$$

$$s(t) = \int_0^t \left| f(u) - g(u) \right| du \quad (3)$$

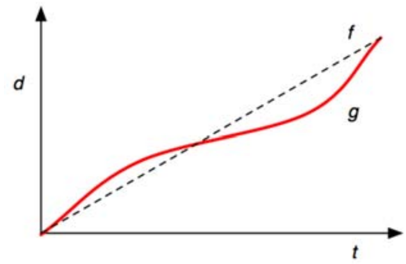

Figure 4. *Graph of the displacement curve for a bow stroke segment.*

We have found that the features that are best suited to the line graph sonification are the *bowing displacement curve*, *bowing height* and the *bow-bridge angle*, since these can be represented as a deviation from a base value. These features can naturally be compared to a zero base value, which would represent a) a linear-velocity stroke, b) a bowing with an equal distance between the fingerboard and bridge, or c) bowing parallel to the bridge. Other features such as the *bow position*, bow velocity and acceleration are not referenced against a base value. For these features only one tone is sounded.

5. RESULTS AND DISCUSSION

The sonifications described here have been presented and discussed with string teachers and students in the i-Maestro user group. They have also been tested in an initial validation phase with six university-level music performance students. Further validation is ongoing.

We have received a range of different opinions about the sonifications, which reflects the fact that string pedagogy is extremely subjective, with many different methods and conflicting ideas. Some teachers found that the auditory alerts/notifications were useful and appropriate whilst others thought they were disturbing and unnecessary in the presence of the teacher. In the tests that we conducted, the students were enthusiastic about the auditory notification since it allowed them to monitor aspects of their performance on their own. They also appreciated the sonification of bow stroke changes for studying the timing of bow stroke gestures.

The sonified line graphs generated some interest; however it was clear that studying the performance via an auditory display was not as immediately understandable to the students as viewing a visual display of the same features. We believe that this is due largely to the familiarity of visual displays, and that users could adapt given more time with the system. The students also found that the pitch of the tone used in the line graph sonification often clashed with the tonality of the music being played, which made it difficult to listen to both simultaneously. One possible solution to this problem would be to combine this system with a score following algorithm, so that rather than using a fixed pitch to signify the base line of the graph, the base pitch could be linked to the musical material. This could make a more useable auditory display, since it would not be necessary to listen to the audio recording and sonification simultaneously.

Some general feedback gathered from the user tests indicates that the sonification and the 3D Augmented Mirror system in general could be beneficial for use with certain students at certain times, and more relevant to advanced performers wishing to sharpen their playing technique, than to beginner students. It was suggested that sonifications would be most appropriate with specially designed bowing exercises rather than for studying pieces. In this way the pitches of the music and the sonification base-pitch can be better related to one another.

We are currently improving our algorithms in order to make more accurate measurements. This work includes detecting exactly which string is being played as opposed to using the current centre-plane approximation. For fast bowing passages it would be desirable to use cameras that were capable of capturing at a higher resolution than 200fps. This would mean that playback rate could be reduced whilst maintaining a high resolution of analysis; however, it could be argued that fluctuations in velocity, acceleration and bow angle have less significance for fast bow strokes, since the tone variation of fast strokes is more difficult for the listener to perceive.

6. CONCLUSION

We presented two types of auditory display designed for use in string instrument training, which have been tested with string teachers and students. Due to the small number of participants in our test, it is clear that further validation is necessary in order to obtain more statistically significant results, and to hear the views of a wider range of people involved in string instrument training. Through our research we found that sonification in real-time, synchronous with a musical performance was most effective with very simple alert sounds. This technique is the least disturbing and therefore would be easiest to integrate into the traditional mode of interaction with in a string lesson. In a non-real-time context it was possible to use sonified line graphs to convey information about the performance. Using an auditory display to study characteristics of a string performance synchronised to an audio recording is a novel concept, which could offer an advantage over a purely visual display due to precise temporal matching.

7. ACKNOWLEDGEMENTS

The i-Maestro project is partially supported by the European Community under the Information Society Technologies (IST) priority of the 6th Framework Programme for R&D (IST-026883, www.i-maestro.org). Thanks to all i-Maestro project partners and participants, for their interest, contributions and collaboration.

8. REFERENCES

[1] Poepel, C and Overholt D. "Recent developments in violin-related digital musical instruments: where are we and where are we going?", in *Proc. of the International Conference on New Interfaces for Musical Expression (NIME 06)*, Paris, France, 2006.

[2] Ng, K., Larkin, O., Koerselman, T., Ong, B., Schwarz, D. Bevilaqua., F. "The 3D Augmented Mirror: Motion Analysis For String Practice Training", in *Proc. of the International Computer Music Conference (ICMC 07),* Copenhagen, Denmark, 2007.

[3] Bevilaqua, F., Guedy, F., Schnell, N., Flety, E., Leyroy, N. "Wireless Sensor Interface and Gesture-Follower for Music Pedagogy", in *Proc. of the International Conference on New Interfaces for Musical Expression (NIME 07),* New York, USA, 2007.

[4] Koerselman, T., Larkin, O., Ng, K. "The MAV Framework: Working with Motion data in Max MSP" in *Proc. of the 3rd International Conference on Automated Production of Cross Media Content for Multi-channel Distribution (AXMEDIS) - i-Maestro Workshop*, Barcelona, Spain, 2007.

[5] N. Rasamimanana, D. Bernardin, M. Wanderley, F. Bevilacqua, "String Bowing Gestures at Varying Bow stroke Frequencies: A Case Study", *Gesture Workshop 2007*, Lisbon, Portugal, 2007.

[6] Ferguson, S. "Learning Musical Instrument Skills Through Interactive Sonification", in *Proc. of the International Conference on New Interfaces for Musical Expression* (NIME 06), Paris, France, 2006.

[7] Verfaille, V. Quek. O, Wanderley M.M. "Sonification of Musicians' Ancillary Gestures", in *Proc. of the International Conference on Auditory Display (ICAD 05)*, London, UK, 2005.

[8] Schnell, N., Borghesi, R., Schwarz, D., Bevilacqua, F., Muller, R. "FTM - Complex Data Structures for Max", in *Proc. of the International Computer Music Conference (ICMC 05)*, Barcelona, Spain, September 2005.

[9] D. L. Mansur, "Graphs In Sound: A Numerical Data Analysis Method for the Blind", M.Sc. Thesis, Department of Computing Science, University of California, 1975.

[10] Brown, L.M. and Brewster, S.A. "Drawing by ear: interpreting sonified line graphs". in *Proc. International Conference on Auditory Display (ICAD 03)*, Boston, USA, 2003.

[11] Brewster, S.A., Wright, P.C. and Edwards, A.D.N. "Parallel Earcons: reducing the length of audio messages". *International Journal of Human-Computer Studies* 43(22):pp. 153-175, 1995.

[12] Neuwirth, E. "Designing a Pleasing Sound Mathematically", *MATHMAG: Mathematics Magazine*, 74, 2001

[13] Brown, L.M., Brewster, S.A., Ramloll, R., Burton, M. and Riedel, B., "Design Guidelines for Audio Presentation of Graphs and Tables", *in Proc. of ICAD 2003 Workshop on Auditory Displays In Assistive Technologies*, Boston, USA, 2003.