The augmented violin project: research, composition and performance report

Frédéric Bevilacqua, Nicolas Rasamimanana, Emmanuel Fléty, Serge Lemouton and Florence Baschet

IRCAM - Centre Pompidou
CNRS STMS UMR 9912
Gesture working group
1 place Igor Stravinsky
75004 Paris – France
+ 33 1 44 78 48 31

{frederic.bevilacqua, nicolas.rasamimanana, emmanuel.flety, lemouton}@ircam.fr, fbaschet@noos.fr

ABSTRACT

In this paper we present the augmented violin developed at IRCAM. This instrument is an acoustic violin with added sensing capabilities to measure the bow acceleration in realtime. We explain first the approach we developed to characterize bowing styles. Second, we describe the realtime implementation of the bowing style recognition system. Finally we describe an electro-acoustic music composition, *Bogenlied*, written for the augmented violin.

Keywords

Augmented violin, hyper-instrument, bowing styles, mapping.

1. INTRODUCTION

The augmented violin project started in 2003 at IRCAM following the interest of several composers to use violin gestures for the control of electronic processes. This project triggered the creation of an interdisciplinary working group on gesture analysis and musical interfaces. One research goal of this group is to work on the concept of "augmented instruments", i.e. acoustic instruments with added gesture sensing capabilities, which is similar to the *hyperinstruments* pioneered at MIT. We believe that such an approach is particularly fruitful for both fundamental gesture research and artistic endeavors.

Different approaches are possible with "augmented instruments". First, sensors can be utilized to add control possibilities that are not directly related to normal playing techniques. For example, various buttons can be added to the body of the instruments. In such a case, the use of sensors implies new gestures for the player. Second, sensors can be applied to capture normal playing gestures. This paper is related to this type of approach, which poses a fundamental question: to what extend can an instrumental gesture, mastered in a particular context, be used in another (or larger) context of musical expression? Such a questioning is actually very fruitful, and remains valid for any types of musical interfaces.

This paper focuses on the ongoing "augmented violin" project. We report here first a study on bowing styles characterization, and the implementation of a system performing realtime bowing style recognition. Second, we describe a music composition, *BogenLied* by Florence Baschet[1] written for the augmented violin and utilizing the analysis system.

2. RELATED WORKS

Different systems have been developed to directly or indirectly measure the violin "gesture parameters". Several systems have been developed over the years using various sensing systems [4][10][9][3]. An alternative approach is to use audio features as a trace of the performed gestures [6][2].

In most of the works, gesture data has been used as a direct control of sound filters, or as the input for physical model synthesis [8]. Fewer works have reported on interpreting the gesture data, in order to provide high-level parameters used in the mapping design. Bernd Schoner [7] used the gesture input to statistically estimate the corresponding sound features in order to drive more expressive synthesis. Chad Peiper [5] used decision trees to classify different bowing styles.

3. APPROACH

Augmented instruments are based on traditional instruments, which gestures are *a priori* defined. In the case of the violin, the different types of bow strokes form a widely accepted and formalized set of gestures. Composition for strings includes bow strokes indications such as *détaché*, *martelé*, *ponticello,* etc. On a finer level, the mastering of various bowing articulations is part of the player's skill.

Our approach has been to use this gesture "vocabulary" as a starting point to build the interaction between the player and the electronics. This is somewhat similar to the approaches of Schoner or Peiper, since we propose to build an "interpretation level" from the data stream, in order to facilitate the mapping between gestures and sounds.

Our first goal was hence to study the relationships between bowing styles and the various sensors data. Such studies allowed us then to build real-time systems that can "interpret"

NIME 06, June 4-8, 2006, Paris, France.

---

[1] Performed at Dijon, France, Nov. 26 2005 Whynote Festival (Premiere), Violin: Anne Mercier (Ensemble Itinéraire); Jan. 07 2006 Maison de la Radio, Paris, France.

the low-level captured parameters into high-level parameters related to bowing styles.

This approach was found in resonance with the compositional approach of Florence Baschet: “With *BogenLied*, my aim as a composer was to focus my attention to the fine instrumentalist gestures, and to use such gesture articulations as the input of an interactive system. Precisely, *BogenLied* is an attempt to create a mixed –acoustic and electronic- sound space, resulting from a sensitive interactive relationship between the soloist and electro-acoustic system. Such interaction would be ideally similar to the type of close relationship between two musicians, as typically found in chamber music”.

4. THE AUGMENTED VIOLIN

Two prototypes have been successively built. They are described below.

4.1 Prototype I

Our first prototype of the wireless bow measurement system is composed of a small electronic board with a microcontroller, two ADXL202 accelerometers from Analog Devices and a digital radio transmitter (fig. 1). The principle of the measurements is similar to the techniques developed by Joe Paradiso and Diana Young [4][10]. Our first prototype utilizes a special radio transmitter that enables collision detection on the carrier and therefore permits to share the transmission bandwidth with other bows equipped similarly. The bow module was very satisfactory. The only drawback is its relative big size due to the thickness of the radio transmitter and the batteries.

4.1.1 Bow acceleration

Accelerations are not analogously sampled but measured using PWM capture with accurate counters and timers. This technique delivers acceleration values on a range of 26000 points, which is significantly better than the use of the internal 10 bit ADC of the microcontroller. A wireless receiver is placed on an Ethersense [1] daughter board. The Ethersense sends the digital data from the accelerometers (as well as others signals) to a host computer through Open Sound Control (OSC).

4.1.2 Bow position

The bow position extraction is directly inspired from the electric field sensing measurement described by Paradiso: two different signals are emitted from the bow tip and frog. A capacitive coupling plate placed behind the bridge collects the mix of the two signals, “tip” and “frog”, which relative intensities depend on the bow position [4]. An Ethersense [1] daughter board demodulates the mix of bow positioning signals. The material used to make the resistive strip placed on the bow was taken from a S-VHS tape. This tape features a homogenously distributed electrical resistance, over its whole length. Two strips of the tape were glued one on the other to protect the resistive side and to make the overall resistance in a more adequate range.

A software calibration was developed to compute the bow position based on the measurement of the two signals "tip" and "frog". The calibration was based on careful measurements for a large set of different bow positions. However, our effort to perform accurate and reliable bow position measurement was deceptive. The main difficulty with this technique is due to the fact that the two measured signals are not sufficient to determine without ambiguity the bow position. This problem especially occurs at the bow extremities where the surface of the "bow plate" diminishes. In addition, the right hand causes a drastic modification of the signals as the impedance of the body interferes with the system, causing a decrease in the signals intensities. An additional difficulty comes from bowing techniques implying significant variations of bow angle, which affects the coupling capacity between the plates.

Overall, the distance measurement was found useful only for qualitative measurement, but problematic for accurate position measurement. Other methods for position measurement are currently experimented.

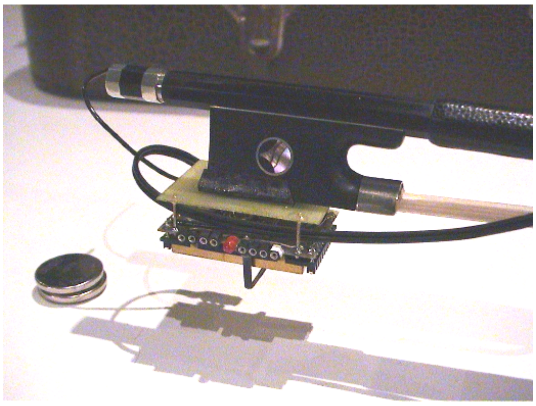

Figure 1: Prototype I.

4.2 Prototype II

A second prototype was built using a subminiature (but single frequency) radio transmitter. The custom-made batteries holder and the second accelerometer were mounted on an extension of the PCB placed on the left side of the bow frog (fig. 2). This second prototype is smaller and lighter than the prototype I (total weight of 17g). The overall thickness of this device was significantly reduced, limiting therefore the risk of scratching the violin. Soft foam was added to totally suppress any possibility of such an accident (not shown on fig 2). Moreover, it consumes less current, allowing 1h30 of continuous playing.

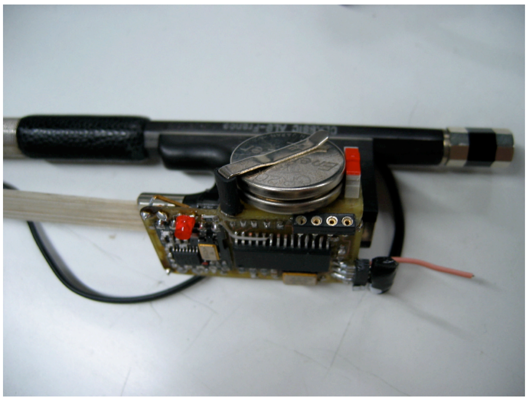

Figure 2: prototype II

5. BOW STROKE STUDIES

We focused our first studies on the following bowing styles: *détaché, martelé* and *spiccato*. After recording these bow strokes in various musical contexts (scales, musical phrases) and with different players, a complete offline analysis has been performed. This analysis has been reported in reference [11] and we summarize here only the important points.

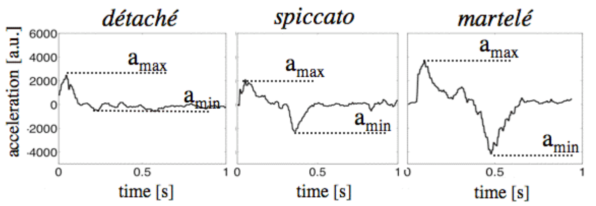

Figure 3. Typical acceleration curves for *détaché*, *martelé* and *spiccato* (from left to right respectively)

- The acceleration curve in the bowing direction shows typically two principal peaks, one positive, a_{max} and the second one negative, a_{min} (for upbow, see fig.3). The *détaché* stroke can be seen as a degenerated case where the second peak is close to zero.

- The use of two parameters a_{max} and a_{min} allows for the clustering of these bowing styles. Recognition rate using the standard *knn* method (k-nearest neighbors) leads to high recognition rates.

- The parameter $(a_{max})^2 + (a_{min})^2$ can be interpreted as a bow stroke "intensity". This parameter is correlated to the energy given to the bow by the musician. It is also correlated to dynamics when the bow is in contact with the strings. For bouncing bowing styles, such as *spiccato*, this parameter gives an indication of the gesture "intensity" rather than the loudness of the sound.

The first study was performed with two violin players, amateur and professional, and showed very consistent results. A second study was performed with a class of 12 students. The results demonstrated than the characterization we proposed remained valid over this larger set of players: the three bow strokes can still be clustered with a_{min} and a_{max}. However, idiosyncratic behaviors were also found, showing that a universal calibration might not be reliable. Interestingly, these results also showed us that this type of analysis could potentially be useful for pedagogical applications. For example, the bowing characterization could provide the students with a complementary feedback to improve bow regularity and/or dynamics. Such potential applications for pedagogy are currently studied.

6. REALTIME BOWSTROKE ANALYSIS

A real time implementation of the bow stroke characterization described in the previous section was implemented in Max/MSP, using several objects of the library MnM [12], dedicated to gesture analysis. The different steps of the gesture analysis are described in Fig 4.

A median filter is used on the x-acceleration data (main bow axis). The dataflow is then split into two different processes run in parallel.

First, a parameter related to the "intensity" of the bowing is computed from the acceleration curve. Precisely, the maximum of the absolute value of the acceleration, computed on a sliding window, is output.

Second, a distinct process allows for the segmentation and characterization of the bow strokes. The varying baseline due to the coupling between angle and acceleration is first removed, (assuming a linear offset). The segmentation is performed then in two steps. First, accelerometer peaks are determined on a sliding window. Second, a procedure sorts and labels the various peaks. One of the difficulties resides in the fact that the "maximum" acceleration a_{max} alternates between bow changes: in upbows a_{max} appears as a local maximum in the raw signal while it appears as a local minimum for downbows. Our selection and labeling procedure is based on the timing between the different peaks. Precisely, we used the fact that the time difference between a_{max} and a_{min} is typically smaller than 100 ms.

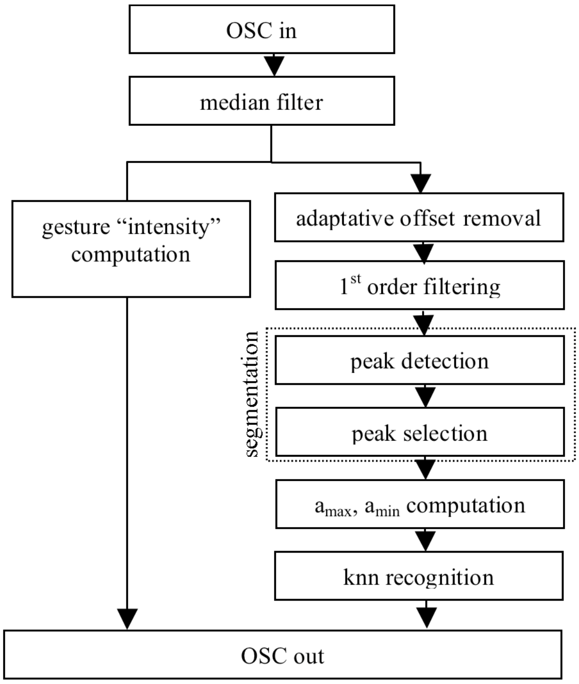

Figure 4. Schematic of the realtime gesture processing

This segmenting procedure is fundamentally underdetermined for a sequence of fast notes, such as a tremolo (when using the accelerometer signal only). In such a case, the peak labeling procedure is ambiguous; different bow strokes might give rise to very similar acceleration curves. Such a problem could be solved by the combination of both audio and acceleration data.

Once the segmentation is performed, the associated a_{max} and a_{min} parameters of each bow stroke are computed. Fig 5 shows an example (as displayed in Max/MSP) of clustered points obtained in the a_{max} and a_{min} plane, with *détaché, martelé* and *spiccato* (playing scales).

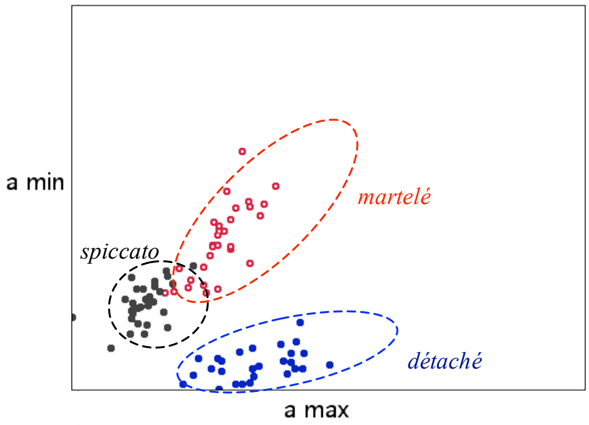

Figure 5. LCD object from Max/MSP showing the clustering on the three bow strokes (real-time): *détaché* (blue), *martelé* (red), *spiccato* (black).

Finally, a *knn* algorithm (k-nearest neighbors) is run to classify each bow stroke as *détaché*, *martelé* or *spiccato*, according to a previously recorded bow strokes database. Weights related to bowing styles are computed using the number of nearest neighbors of each class. The weight averages over several notes are also computed.

The final results, i.e. gesture intensity and bow stroke weights are sent through Open Sound Control (OSC) to the computer controlling the sound processes.

7. COMPOSITION WORK: *BOGENLIED*

7.1 Setup

The diagram (figure 6) shows the configuration used in the performance setting. Two separate computers (two Macintosh G4 PowerBooks) are running Max/MSP, and communicating using OSC. The "sound processing" computer generates the digital sound environment using the live violin sound, captured by microphone. Various parameters of the sound processing are controlled by the gesture data transmitted from the "gesture-processing" computer. The electronic sound is spatialized (using the "Ircam spat") and rendered by a hexaphonic sound diffusion system.

7.2 *Bogenlied* form.

As shown in fig. 7, the musical form of BogenLied is a simple linear form divided into nine sections, alternatively with and without electronics. In section *II*, *IV* and *VI*, the musical writing is focused on specific articulations of bow strokes: *détaché*, *martelé*, *spiccato*. Each of these sections is associated to specific compositional materials, as well as to specific electronic sound processes (described in section 7.3).

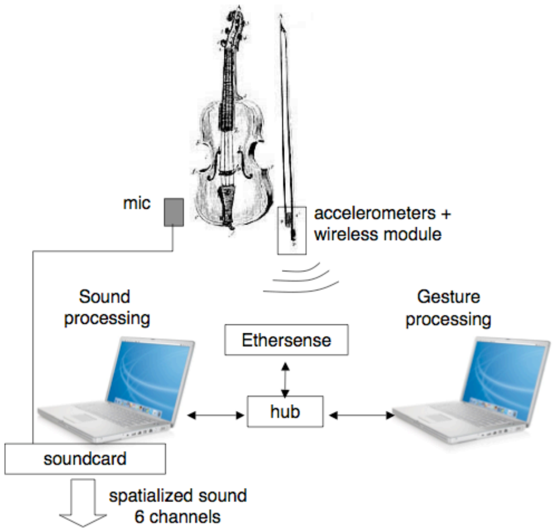

Figure 6. Setup for the piece *Bogenlied*

In section *VIII*, the three bowing styles appear successively, along with their associated compositional material. In this section, the choice of the electronic processes is driven by the recognition system.

In last section (*IX*), the performer plays hybrid bow strokes, which sequence appears as a series of "mutations" from one articulation to another one. As described further below, the electronic sound is then built as a combination of the electronic processes performed in section *II*, *IV* and *VI*.

section	bowing style	sound	Gesture analysis
I		acoustic	
II	*détaché*	mixed	gesture intensity
III		acoustic	
IV	*martelé*	mixed	gesture intensity
V		acoustic	
VI	*spiccato*	mixed	gesture intensity
VII		acoustic	
VIII	*all 3 bowing styles*	mixed	gesture intensity + bow stroke recognition
IX	*hybrid*	mixed	gesture intensity + bow stroke recognition

Figure 7. Structure of *Bogenlied*

7.3 Audio processing

The Max/MSP sound patch contains two granular synthesis modules and several effects. The first granulator is used to produce drones (pedal-notes) triggered by the bow strokes. A separate drone is linked to each bowing styles: G, A, and E for *détaché*, *martelé* and *spiccato*, respectively.

The second granulator is used for real-time processing of the violin sound. The audio signal from the microphone is stored in a 5 second circular buffer~, and 100 to 200 ms grains are played back from this buffer. The granular synthesis sound is processed through several standard effects: frequency-shifter, harmonizer, filter and vocoder. Different combinations and parameterizations of these effects are used in each section. Some parameters are controlled in realtime by the gesture data, as explained next.

7.4 Gesture Mapping

The mapping between the gesture data and sound processing parameters is built during the piece as a gradual superimposition of three different mapping modalities:

a) *Continuous* mapping, applied indifferently to all bow strokes

- sections *II* , *VI*: the gesture "intensity" is linearly mapped to the grain density (number of grains played simultaneously)
- section *IV:* the gesture "intensity" is mapped to the grain transposition spread (small gestures induce small or no transposition, strong gestures cause all the grains to be randomly transposed around the played pitch).

b) S*elective* mapping using bow stroke recognition:

- sections *VIII* and *IX*: bow stroke recognition is used to select the type of sound processing. The recognition of a *détaché, martelé or spiccato* recalls the presets of section *II*, *IV* or *VI,* respectively (including the corresponding drones).

c) Mapping *mixing*, using bow stroke characterization (of hybrid bow strokes):

- section *IX*, the performer plays hybrid bow strokes. The gesture analysis computes for a given bowstroke three weights, corresponding to "likelihood" to be related to the different bowing styles (in other words a given bow stroke is considered as a mix between different bowing styles). These weights are then used to control the mixing levels of the drones associated with each bowing style. Each bow stroke is thus colored by the combination of these three pitches reflecting continuously the quality of the articulation.

8. DISCUSSION AND PERSPECTIVES

We report here developments related to our current "augmented violin" project. From the technological point of view, our prototypes were tested on several experiments and performances and found to be robust. The added weight on the bow seems manageable for the players: the various professional violinists we worked with agreed to play with such a constraint.

We described the real-time implementation of a bow stroke analysis framework that was reported previously [11]. This approach requires the segmentation of the acceleration data stream in separate bow strokes. Such a task is difficult when the acceleration signal alone is used, due to various artifacts. The algorithm we designed was generally satisfactory, but was not applicable to fast notes, such as tremolo. We are currently investigating other approaches, based for example on Hidden Markov Models, in order to overcome such limitations.

BogenLied is the first piece written for the IRCAM augmented violin. It is worth to note that this piece (composed by Florence Baschet) has been developed in the context of an interdisciplinary workgroup. The piece takes advantage of the gesture analysis we described. In particular, the mapping is expressed from elements of the musical language, in order to create an electronic environment that performers can apprehend intuitively. Our collaboration with the violinist Anne Mercier confirmed us that such an approach is very effective.

The real-time analysis was reliably used in performances of *BogenLied*. Overall, the recognition system was satisfactory. In particular, the performer was able to control the mapping easily with her own gestures, and she could feel that the system was reacting well to her gesture. A promising point was the possibility to characterize "hybrid" bow strokes by continuous parameters. This feature seems to offer pertinent information in complex musical phrases.

In the near future, other performers will experiment with this system, and we are expecting interesting comparisons. In particular, the modules will be used with other string instruments such as the viola, cello and double bass. This collaborative work has provoked a high interest at IRCAM and several other artistic works are currently in progress.

9. ACKNOWLEDGMENTS

We would like to thank Anne Mercier, Alain Terrier and Jean-Loup Graton for their precious contribution and help.

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.net, www.i-maestro.org). Thanks to all I-MAESTRO project partners and participants, for their interests, contributions and collaborations.

10. REFERENCES

[1] E. Fléty, N. Leroy, J.-C. Ravarini, and F. Bevilacqua. Versatile sensor acquisition system utilizing network technology. *In Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2004.

[2] C. Goudeseune, G. Garnett, and T. Johnson. An interface for real-time classification of articulations produced by violin bowing. *In Proceedings of the Conference on Human Factors in Computing Systems*, 2001.

[3] D. Overholt. The overtone violin. *In Proceedings of the Conference on New Instruments for Musical Expression, NIME,* 2005.

[4] J. A. Paradiso and N. Gershenfeld. Musical applications of electric field sensing. *Computer Music Journal*, 21(2), 1997.

[5] C. Peiper, D. Warden, and G. Garnett. An interface for real-time classification of articulations produced by violin bowing. *In Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2003.

[6] C. Poepel. Synthesized strings for string players. *In Proceedings of the Conference on New Instruments for Musical Expression, NIME*, 2004.

[7] B. Schoner, C. Cooper, C. Douglas, and N. Gershenfeld. Cluster-weighted sampling for synthesis and cross-synthesis of violin family instrument. *In Proceedings of the International Computer Music Conference, ICMC*, 2000.

[8] S. Serafin and D. Young. Bowed string physical model validation through use of a bow controller and examination of bow strokes. *In Proceedings of the Stockholm Music Acoustics Conference, SMAC*, 2003.

[9] D. Trueman and P. R. Cook. Bossa: The deconstructed violin reconstructed. *In Proceedings of the International Computer Music Conference, ICMC,* 1999.

[10] D. Young. The hyperbow controller: Real-time dynamics measurement of violin performance. *In Proceedings of the Conference on New Instruments for Musical Expression, NIME*, Dublin, Ireland, 2002.

[11] N. Rasamimanana, E. Fléty, F. Bevilacqua, Gesture Analysis of Violin Bow Strokes, GW 2005, Lecture Notes in Artificial Intelligence 3881, S. Gibet, N. Courty, and J.-F. Kamp (Eds.): pp. 145–155, 2006.

[12] F. Bevilacqua, R. Muller, N. Schnell, MnM: a Max/MSP mapping toolbox. *In Proceedings of the 2005 International Conference on New Interfaces for Musical Expression, NIME*, Vancouver, p. 85-88, 2005