Extracting Human Expression For Interactive Composition with the Augmented Violin

Mari Kimura
The Juilliard School
60 Lincoln Center Plaza, New York, NY 10023 USA
marikimura@me.com

Nicolas Rasamimanana
Phonotonic
84 rue du Faubourg St Martin, 75010 Paris, FRANCE
IRCAM
1, place Igor-Stravinsky, 75004 Paris, FRANCE
nicolas@phonotonic.net

Frédéric Bevilacqua
Bruno Zamborlin
Norbert Schnell
Emmanuel Fléty
IRCAM
1, place Igor-Stravinsky, 75004 Paris, FRANCE
Frederic.Bevilacqua@ircam.fr

ABSTRACT
As a 2010 Artist in Residence in Musical Research at IRCAM, Mari Kimura used the Augmented Violin to develop new compositional approaches, and new ways of creating interactive performances [1]. She contributed her empirical and historical knowledge of violin bowing technique, working with the Real Time Musical Interactions Team at IRCAM. Thanks to this residency, her ongoing long-distance collaboration with the team since 2007 dramatically accelerated, and led to solving several compositional and calibration issues of the Gesture Follower (GF) [2]. Kimura was also the first artist to develop projects between the two teams at IRCAM, using OMAX (Musical Representation Team) with GF. In the past year, the performance with Augmented Violin has been expanded in larger scale interactive audio/visual projects as well. In this paper, we report on the various techniques developed for the Augmented Violin and compositions by Kimura using them, offering specific examples and scores.

Keywords
Augmented Violin, Gesture Follower, Interactive Performance

1. INTRODUCTION
The Augmented Violin project brings together the development of motion capture systems and the study of bowed string performance. The outcome of this project is essentially an "augmented bow", which can be used to control digital media. It is not an extra contraption one needs to install on the acoustic violin or bow, nor a mechanical device to replace the existing instrument. The Augmented Violin honors the violin as is, and simply extracts expression from what is already happening in performance, which seems like the most practical approach that makes sense for musicians.

To start our research, we first laid out specific goals related to the potential compositional schemes pursuing our work with IRCAM's Augmented Violin (customized) and using a new series of software tools, such as the gesture recognition system based on the MuBu-GF platform [3] as well as other specific developments.

NIME'11, May 21-23, 2011, University of Michigan, Ann Arbor.

Specifically, these goals were: 1) develop the "Clone" performance, i.e. controlling through bowing gestures the speed of live recorded material; 2) Recognizing different phrases apart in real time; 3) Using Gesture Following (GF) to interact with pre-recorded sounds in real time; and 4) Analyzing and interacting with "anticipatory" or "before and after" gestures of bowing. Several effective results were achieved, which became very useful tools in interactive performance for the violin.

2. BACKGROUND
At NIME 2006 conference, Kimura learned about IRCAM's Augmented Violin, and since has been working on creating several compositions and presenting more than dozen premieres and performances incorporating the Augmented Violin into her compositions. Improving on the original design of the Augmented Violin, which required wearing a wrist band to house the battery portion connecting to the sensor attached to the bow with an 'umbilical' cord, Kimura commissioned a custom-fit glove created by New York fashion designer Mark Salinas. This glove, worn on the right hand, houses both the sensor and battery portions of the Mini-MO, a wireless 3D accelerometer and 3-axis gyroscope motion sensor developed by Emmanuel Fléty [17]. The Mini-MO communicates with the computer via an Ethernet receiver, thus freeing the violinist's bow arm completely from any wiring. The thin, flexible latex fabric of the glove houses the small device un-intrusively for performance: a practical solution for violinists (see Figure 1).

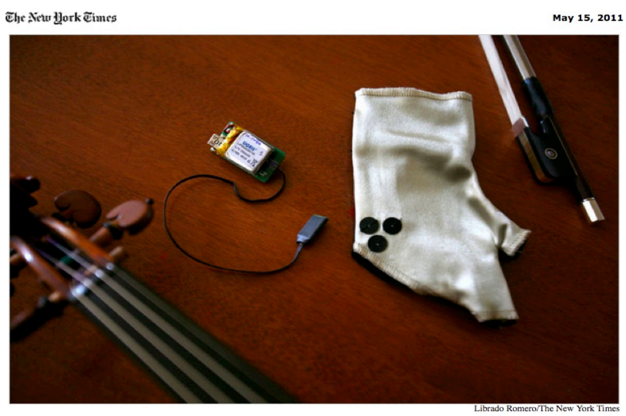

Figure 1. The Augmented Violin Glove, with the latest Mini-MO (Modular Musical Objects) created by IRCAM.

In 2007, during her residency at Harvestworks in NYC, Kimura created her first work for Augmented Violin entitled *VITTESSIMO* [4] in collaboration with Frédéric Bevilacqua; she also described the creative process of this work [5]. For a general review of related works on augmented and hyper-instruments based on strings instruments, we refer the reader to Poepel, Overholt [15] and Bevilacqua *et al* [16].

3. RESEARCH

3.1 'Cloning' with MuBu-GF

CLONE BARCAROLLE (2009) by Kimura [4] uses the 'cloning' function with GF. In this work, the movement of the bow is 'cloned' together with the audio data. When a violinist repeats the bowing movement, the computer recreates the performance, thus enabling the violinist to play with a 'clone' that accompanies him/her along according to the bowing movement. In *CLONE BARCAROLLE,* the computer 'clones' the ostinato, creating a virtual duo with the violinist. For a performer, it is quite a departure from playing along with a simple delay, which is an exact 'echo' or a 'repeat' of one's performance. 'Cloning' is a new musical language for musicians; it can have a 'life' of its own, capable of being controlled after being recorded. It requires a new way of processing musical information during performance.

The "cloning" is made possible using the gesture follower (using the *gf* and *mubu* externals in Max). When a musical phrase is played for the first time, the patch records both the audio and the sensor parameters. The audio is stored in a buffer, while also analyzed to obtain the audio energy profiles. These audio energy profiles are then combined to the sensors parameters of Mini-MO to form a multimodal set of data that is memorized by the gesture follower. All parameter profiles are normalized (between 0 and 1) and form a complementary set of descriptors of the phrase: the accelerometer/gyroscope give very precise information on the bowing attacks and releases while the audio energy gives more global information on the phrase morphology.

When the same phrase is played the second time, the system is set to "follow mode". In this case, the multimodal parameters (audio energy and sensors values), measured in real time, are compared to the stored data. This allows for the continuous estimation of playing speed. The original buffer can thus be replayed using the current playing speed. For this, a phased vocoder is used to preserve the original speed (superVP). The replay of the audio buffer is thus the "clone", precisely synchronized to the current playing. This adaptive behavior provides the musician with a range of expressivity different from the standard use of "record/play" of samples with a fixed speed. (See Figure 2. Excerpt from "Clone Barcarolle" (2009))

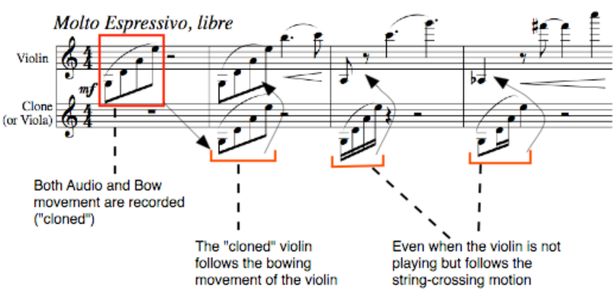

Figure 2. Excerpt from "Clone Barcarolle" (2009).

3.2 Recognizing phrases in real time

Many musical scenarios and trials were made, recording and analyzing multiple musical phrases using a "multi-record" function (Figure 3), recording short sample compositions including audio, video and bowing motion all together, synchronized for analysis (Figure 4).

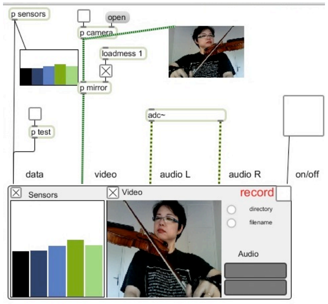

Figure 3. "multi-record" MaxMSP patch, recording audio and video, and bowing gesture in combination.

Figure 4. Phrase-recognition test phrases: areas in gray squares tested for tracking by GF.

After these many phrase-recognition tests, we decided it is better to segment and concatenate musical material than trying to build a system that recognize phrases, since the purpose of this research is not to create a "score follower". In order to extract musical expression so as to interact with the computer in real-time performance, we found that GF works best when tracking short and distinct bowing motions.

Various short gestures such as *pizzicato, tremolo, bowing direction, détaché, spiccato* etc. are now possible to detect using the Augmented Violin.

3.3 Analyzing "before and after" gestures

We have been interested in separating two kinds of bowing motions: 1) bowing motions that actually make sounds, or "functional" bowing movement, and 2) "anticipatory" or "before and after" movements which violinists make before and after the functional strokes. The latter movements are made in order to prepare or conclude the musical expression before or after the "functional" sound-making motions are made.

Any string player knows there is much expressive information in bowing motion in these "before and after" moments. We conducted experiments by playing the same short phrase and applying different expressions, and recorded them using the aforementioned "multi-record" function. A melody is played with different musical expressions: 1) “dolce” throughout, ending softly; 2) “dolce” making a sudden shift to “agitato” at the end and ending loudly (see Figure 5). We tracked the bowing motion focusing especially on the portions highlighted in gray strips. This is when the violinist is not making any sound, but concluding one expression and preparing the bow to perform the next. This experiment led us to develop several tools for tracking musical expression from the bowing movements described below.

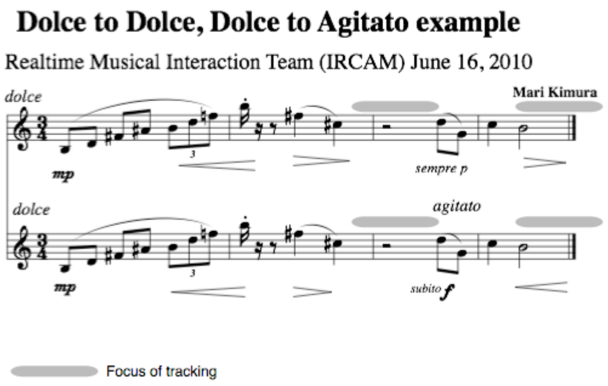

Figure 5. “Before and After” stroke analysis example

3.4 Soft / Loud Ending

Kimura’s pieces are based on three main bowing gesture processing. They were designed upon the concept of movement qualities [6], where the goal is not to achieve a particular gesture or action but rather a way of performing it, i.e. a quality. We believe such an approach to be promising as it can provide for a higher level of abstraction in interaction design, while still keeping a sense of mediation [7].

Ending a performance can be done in many different manners. We decided to try detecting musical expressions including two different musical endings, shown in this example: 1) soft ending, 2) loud ending. Using the results of the analysis, we developed a simple “ending detector”, analyzing the combination of physical and musical gestures: right hand movement together with sound and musical context for the violin. (see Figure 6).

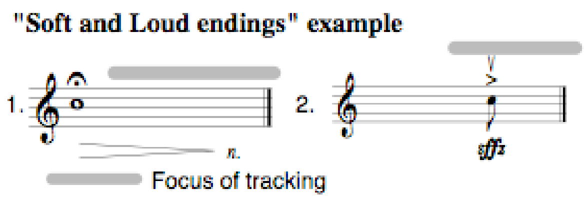

Figure 6. Soft (1) and Loud (2) ending tracking.

We here wanted to focus on soft on-the-string endings, versus loud off-the-string endings. The Soft/Loud Ending detection was designed to grab such ending qualities of Kimura’s performance. This detector is combining gesture and audio cues. It simply relies on bow orientation (tilt) when sound is fading away (audio energy). Based on this, two different configurations are registered, soft ending and loud ending. Through a k-Nearest-Neighbor algorithm operating on bow tilt and audio energy, we are able to distinguish in real time the performed ending quality.

3.5 “Stillness” detector

With a realization that the cornerstone of a ‘bowed’ instrument is the ability to sustain sound, unlike ‘plucked’ instruments, we wanted to simply detect if and when the bow has achieved a “steady” state while holding a note. This “stillness” detector became one of the most important interactive methods in Kimura’s compositions.

The stillness detector actually tracks two kinds of bowing movements: smooth sustained movements and jerky fast movements. As such, it can be related to the musical flow created by the artist through the alternation of short and long, energetic and smooth strokes.

This gesture descriptor is based on the acceleration in the bowing direction. It enables one (1) to detect bow sustained movements and (2) to register the time this movement was maintained. ‘Stillness Detection’ can be used to recognize the sustained characteristics of a bowing movement, i.e. one of its movement qualities. Registered time can be used to differentiate short and long strokes, e.g. by setting threshold values, or to design time evolving interactions directly depending on the parameter value.

To compute the stillness gesture descriptor, we first derive acceleration in the bowing direction to get the movement smoothness, aka ‘jerk’. This value accounts for the level of movement smoothness. From this value we are able to distinguish between two states. ‘Still’ corresponds to low jerk values, i.e. smoothest movements, as opposed to ‘Non-Still’, with higher jerk values. This detection can be stabilized using an additional threshold for minimum time in a state. These two threshold values actually depend on each violinist, each musician having his/her own movement characteristics, and can be set during practice. The “stillness” detection is prominently demonstrated in Kimura’s work “*Pédalez, Pédalez*!” (2010). Using the “stillness” detection Kimura created a sustain pedal without using a foot pedal commonly used by electric violinists. (see Figure 7).

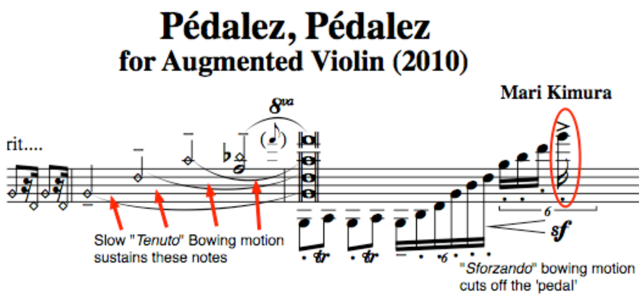

Figure 7. “Pédalez, pédalez!” excerpt, creating a sustain pedal without a foot pedal.

This “stillness” detector turned out also as a useful substitute for the *superVP.ring~* external in Max. A ring buffer is rather cumbersome to control in real time performance, since it constantly records a certain amount of the “past” then “rolls off”. Using the “stillness” detector, it is possible to simulate the ring buffer, which is used in Kimura’s composition *CANON ELASTIQUE* (2010). Using two *superVP.scrub~* externals controlled by “stillness” detection, it was possible to constantly change the 2nd voice (the “past”) by applying “stillness” control so as to follow the 1st voice (the performer). In *CANON ELASTIQUE,* one can create an “elastic” past (see Figure 8).

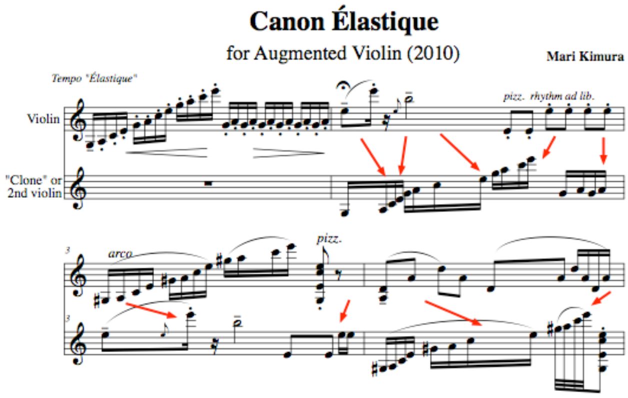

Figure 8. "*Canon Élastique*" creating 'elastic past'.

3.6 Agogic player, "flow-follower"

Another application of "stillness" detection is "agogic player", or what Kimura calls a "flow-follower", an alternative to a "score follower". Performers in real life don't have to follow note by note, beat by beat, in order to play together in sync. Rather, they follow each other's musical "flow" together, or agree on musical "common sense" of the particular musical scenario, a piece of music or context. By using "stillness" detection, it is possible to demonstrate how a quantized piano performance (either MIDI or audio) can co-perform, or follow the violinist's musical flow without using a score follower. Even when the violinist is preparing a stroke without playing, the piano part can naturally and musically be slowed down before the violin's entrance (see Figure 9).

In this short excerpt of a Bossa Nova tune by Tom Jobim, which requires lucid tempo and *rubato,* the violinist starts playing as the piano plays the 8^{th} note scale upward. The violinist' right arm movement at this point is still preparing to play the first note, and it is 'steady'. By applying the "stillness" detection here, it is possible to naturally slow down the piano part slightly, as any sensible human pianist would in real life.

In the second measure, the "stillness" detector also applies "agogic" or a small 'breath' between the beats. As a result, the computer piano naturally plays along, following the flow of the music created by the bowing, making human and machine interaction mutual and omni-directional: not one triggering and the other obeying.

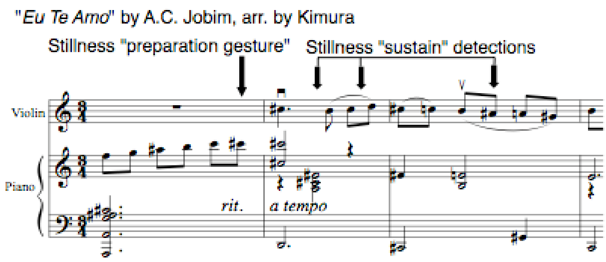

Figure 9. "Flow Following" between human and computer.

3.7 Combining OMAX with Gesture Follower

While working on concatenating musical material using GF, we found interesting to join this system with the improvisational system called OMAX, developed by the IRCAM's Musical Representation Team [8,11]. OMAX, which runs on MaxMSP, creates a "co-improvisation" that learns the player's improvisation performance and plays back the concatenated performance in real time. OMAX was mainly designed to include a "computer operator" who controls its behavior such as 'fading out' to end its improvisation.

Using GF interacting with OMAX, it is possible to create an autonomous improvisation without the computer operator performing OMAX. Working with the distinct behavior of OMAX, and with the aforementioned various gesture detections including "ending" or "stillness" detector, it is now possible to terminate OMAX's improvisation from pre-defined concatenated gestures, such as pizzicato or sustained fading tones. Kimura's compositions using OMAX and GF, to date, include *VIOMAX (2010)* and *Duet x2* for violin and cello (2011) using two Mini-MOs (see Figure 10).

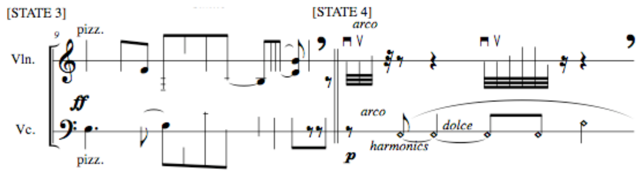

Figure 10. Excerpt from "*Duet x2*": 'stillness', 'pizzicato', 'tremollo' concatenated gestures used to control OMAX.

Technically the same system is used for the *VIOMAX* piece as in *CLONE BARCAROLLE*, but taking advantage of different features of the gesture follower. First, the data is not recorded during the performance but during rehearsals: the multimodal data is recorded to disk and read before the performance. Second, the real time data is not compared to a single recorded musical phrase as in "cloning", but to several phrases. In this case, the system estimates which one of the phrases is the most similar to what is currently played: this mode is called the recognition mode. Importantly, the system can be easily loaded with different phrases to be recognized, which just need to be recorded once by the musician. This flexibility is important to adapt to a vocabulary that can designed by the musician and potentially evolve over time.

4. CURRENT AND ONGOING

4.1 Recent multi-media projects

The Augmented Violin can be used for audio/visual projects as well as collaborations. Kimura's work *PHANTOM* (2009) is a collaborative project with New York-based artist Toni Dove in her multimedia work, "*Lucid Possessions*". In *PHANTOM*, the Augmented Violin controls interactive video and a robot with 3D projections.

More recent audiovisual works with the Augmented Violin include "*VOYAGE APOLLONIAN*" for interactive audio/video (visuals created by graphic artist Ken Perlin) [10]. The Augmented Violin was used to control the Quicktime movie using "stillness" and other concatenated gesture detections.

4.2 Eigenspace

Kimura's most recent project is entitled *EIGENSPACE*. For violinists, bowing motion is kind of an 'arc', which involves arm motions that include 'inertia' movements before and after the strokes. In *EIGENSPACE (2011)*, this 'inertia' movement is incorporated to interact with the audio and video.

We came up with a concept to simulate this motion with a spring, using "eigenvalues". This last gesture processing introduces a physical behavior between bow gesture data and sound (or video) parameters. Similarly to works by Henry [13], interaction here is grounded on a damped mass spring system. In our case, however, the physical system is controlled by the bowing energy computed from bow acceleration. The musician can therefore imprint different movements to the physical system, which reacts according to its settings. Changing its parameters, like weight, mass or stiffness factor, alters the system reaction depending on the created eigenmodes. These eigenmodes can be related to different movement qualities as they exhibit various behaviors for a same mass displacement: jerky, smooth, or periodic. Playing on these modes enables to create an organic relation between the artist and the computer.

In *EIGENSPACE*, the graphics created by the team of Japanese movie director Tomoyuki Kato interact with the eigenvalues and the 'spring' motion during the 'retake' of the bow. Extracting the 'retake' motion of the bow, the 'spring' or the inertia mechanism, can be calibrated to the desired 'stiffness', 'reaction time' and 'damping' according to the composer's musical needs and the violinist's expressive bowing motion.

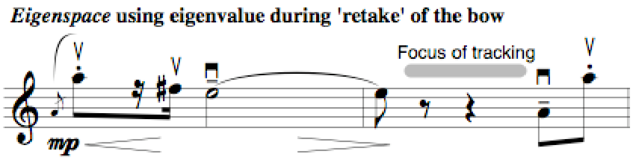

Figure 11. "*Eigenspace*" excerpt: the 'retake' of the bowing motion interacts with interactive graphics.

Figure 12. "Eigenspace" premiere at Roulette in Brooklyn, NY, Oct. 9th, 2011.

5. CONCLUSION

The Augmented Violin is one of the most unintrusive sensor devices available for violinists to provide bowing and motion data to a computer media system such as MaxMSP. Its main appeal to violinists is that it honors and enhances the existing instrument without any wirings and contraptions, which should be more user-friendly to average string players.

In this article, we have presented various compositional techniques that build on the unique capabilities of the Augmented Violin, illustrated through Kimura's compositions. These include the ability to "clone" oneself, while applying various transformations to the "clone" controlled by the player via the Augmented Violin; recognizing different phrases in real time; analyzing "before" and "after" gestures; recognizing "soft" and "loud" endings; detecting "stillness" and following musical flow. We have shown how to combine the gesture follower with other systems, like the improvisational system OMAX (also developed at IRCAM), and into multi-media collaborations projects between Kimura and others. The Augmented Violin has also inspired other composers to create interactive compositions that had been previously impossible.

We will need to further develop user-friendly interfaces, thus affording many string players to try out this elegant musical system of the future.

6. ACKNOWLEDGMENTS

Special thanks to IRCAM's Real Time Musical Interaction Team, and IRCAM's Musical Representation Team, Gérard Assayag and Benjamin Lévy. This work was made possible through the Composer-in-Residence in Musical Research program at IRCAM. Thanks to Hervé Brönnimann for his help in the preparation of this manuscript.

7. REFERENCES

[1] F. Bevilacqua, F. Guédy, N. Schnell, E. Fléty, and N. Leroy. "Wireless sensor interface and gesture-follower for music pedagogy". In *Proceedings of the 2007 Conference on New Interfaces for Musical Expression (NIME07)*, New York, NY, USA, 2007.

[2] F. Bevilacqua, B. Zamborlin, A. Sypniewski, N. Schnell, F. Guédy, and N. Rasamimanana. "Continuous realtime gesture following and recognition". In *Embodied Communication and Human-Computer Interactin*, volume 5934 of Lecture Notes in Computer Science, pages 73–84. Springer Berlin / Heidelberg, 2010.

[3] N. Schnell, A. Röbel, D. Schwarz, G. Peeters, and R. Borghesi. "MuBu & Friends - Assembling Tools for Content Based Real-Time Interactive Audio Processing in Max/MSP". In *Proceedings of the 2009 International Computer Music Conference (ICMC)*, Montreal, Canada, 2009.

[4] M. Kimura. *The World Below G and Beyond*, solo CD. Mutable Music label, 2010.

[5] M. Kimura. "Making of Vitessimo for Augmented Violin: Compositional Process and Performance". In Proceedings of the *2008 Conference on New Interfaces for Musical Expression (NIME08)*, Genova, Italy, 2008.

[6] S. Fdili Alaoui, B. Caramiaux, and M. Serrano. "From dance to touch: movement qualities for interaction design". *CHI*, ACM (2011), 1465–1470.

[7] M. Leman. *Embodied music cognition and mediation technology*. The MIT Press, Cambridge, Massassuchetts, 2007.

[8] G. Assayag, G. Bloch and S. Dubnov. "Introducing Video Features and Spectral Descriptors into the Omax Improvisation System". Demo*, International Computer Music Conference*, Belfast, Ireland, August, 2008.

[9] Assayag, G. , Bloch, G. "Navigating the Oracle: a Heuristic Approach". In *Proceedings of the 2007 International Computer Music Conference (ICMC),* Copenhagen, Denmark, 2007.

[10] A. Cont, S. Dubnov, and G. Assayag. "Anticipatory Model of Musical Style Imitation using Collaborative and Competitive Reinforcement Learning". In *Anticipatory Behavior in Adaptive Learning Systems*, (Martin Butz and Olivier Sigaud and Gianluca Baldassarre, Berlin), 2007.

[11] G. Assayag, G. Bloch, M. Chemillier, A. Cont, and S. Dubnov. "Omax Brothers : a Dynamic Topology of Agents for Improvization Learning". In *Workshop on*

Audio and Music Computing for Multimedia, ACM Multimedia 2006, Santa Barbara, 2006.

[12] L. Greenemeier. “String Theory: Violinist Taps Artificial Intelligence to Interact with Her Unique Sound”. *Scientific American*, May 31, 2011.

[13] C. Henry. “Physical Modeling for Pure Data (pmpd) and real time interation with an audio synthesis”. In *Proceedings of the Sound and Music Computing Conference, SMC,* 2004.

[14] N. Rasamimanana, F. Bevilacqua, N. Schnell *et al.* “Modular Musical Objects Towards Embodied Control Of Digital Music”. *TEI*, (2011), 9-12.

[15] C. Poepel, D. Overholt. “Recent developments in violin-related digital musical instruments: where are we and where are we going?” In *Proceedings of the International Conference on New Interfaces for Musical Expression,* Paris, France, pp. 390–395, 2006.

[16] F. Bevilacqua, F. Baschet, and S. Lemouton. The augmented string quartet: experiments and gesture following, *Journal of New Music Research,* 2012.

[17] E. Fléty, and C. Maestracci. “Latency improvement in sensor wireless transmission using IEEE 802.15.4.” In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 409–412, Oslo, Norway, 2011.