

# ACQUISITION OF VIOLIN INSTRUMENTAL GESTURES USING A COMMERCIAL EMF TRACKING DEVICE

*Esteban Maestre, Jordi Bonada, Merlijn Blaauw*

*Alfonso Pérez, Enric Guaus*

Music Technology Group, Universitat Pompeu Fabra

Ocata 1, Barcelona, SPAIN

{emaestre, jbonada, mblaauw, aperez, eguaus}@iua.upf.edu

## ABSTRACT

This paper presents a method for the acquisition of violin instrumental gesture parameters by using a commercial two-sensor 3D tracking system based on electro-magnetic field (EMF) sensing. The methodology described here is suitable for acquiring instrumental gesture parameters of any bowed-string instrument, and has been devised by paying attention to intrusion, flexibility, and robustness. After reviewing relevant related work in the field, we give an overview of the application context, pointing out some basic needs to be fulfilled for our research purposes. Then, we present the steps for calibrating the system, followed by details on the computation of a number of relevant instrumental gesture parameters. The use of a number of the extracted parameters to perform score-performance alignment and database automatic annotation is also outlined. Finally, we conclude stating next steps in using acquired data, along with further developments of the methodology.

many instrumental gesture parameters as possible, (2) to be accurate, (3) to be easily attached to any violin/bow, and (4) to present a small effect on the instrument playability. As a first step towards our research goals, we use the acquired parameters for automatically annotating a performance database by applying a custom score-performance alignment procedure.

The paper is structured as follows. First, we offer an outlook to the related work in the field, highlighting the main differences with the approach presented here. Then, we put our work into context, pointing out our research purposes, and discussing the fulfillment of our application requirements. Section 4 outlines the calibration procedure, while Section 5 gives the details of the computation of a number of instrumental gesture parameters. In Section 6, we present the score-alignment procedure devised for annotating our database. We conclude by discussing further developments of the methodology and future uses of our gathered data.

## 1. INTRODUCTION

Interaction between performer and instrument during a performance gets more complex when dealing with excitation-continuous musical instruments, as it is the case of bowed-string family, which is often considered among the most articulate and expressive musical instrument family. Capturing violin instrumental gestures parameters (i.e. directly involved in sound production mechanisms) is of great interest, as there have been already a number of works dealing with this issue and with the aim of applying results both to performance and to research (see Section 2). Particularly for our research purposes, focused on the study of instrumental gestures in violin traditional performance in order to apply them for synthesis, it is important to be able to acquire relevant timbre-related parameters with enough accuracy, while not affecting the performance habits.

In this paper, we present a new approach for acquiring violin instrumental gesture parameters in real-time. This approach is based on the use of electro-magnetic field (EMF) sensing (two wired sensors), by using one of the commercially available solutions. The main requirements that we forecasted when devising the proposed methodology were: (1) to have a system capable of gathering as

## 2. RELATED WORK

Research in capturing gestures involved in sound production mechanisms of bowed-string instruments has led to diverse successful approaches that can be found in the literature since already the eighties.

In [1] and [2], Askenfelt presents methods for measuring bow motion and bow force using diverse custom electronic devices attached to both the violin and the bow. The bow transversal position is measured by means of a thin resistance wire inserted among the bow hairs, while for the bow-bridge distance, the strings are electrified, so that the contact position with the resistance wire among the bow hairs is detected. For the bow pressure, four strain gages (two at the tip and two at the frog) are used. A different approach was taken in [6], where Paradiso and Gershenfeld measured bow displacement by means of oscillators driving antennas (electric field sensing). In a first application carried out for cello, a resistive strip attached to the bow was driven by a mounted antenna behind the bridge, resulting as well into a wired bow. Afterward, in the violin implementation of this methodology, which resulted into a first wireless measurement system for bowing parameters, the antenna worked as the re-

ceiver, while two oscillators placed in the bow worked as drivers. What they call bow pressure is measured by using a force-sensitive resistor below the forefinger (or between the bow hair and wood at the tip). These approaches, while providing means of measuring the relevant bowing parameters, did not allow tracking performer movements. Furthermore, the custom electronic devices that needed to be attached to the instrument resulted to be somehow intrusive, while not being easy to interchange the instrument at performer's demand.

More recent implementations of violin bowing parameter measurement introduced some important improvements, resulting into less intrusive systems than previous ones. In [9] and [10], Young measured downward and lateral bow pressure with foil strain gages, while bow position with respect to the bridge is measured in a similar way as in [6]. The strain gages are permanently mounted around the midpoint of the bow stick, and the force data is collected and sent to a remote computer via a wireless transmitter mounted at the frog. Besides the additional hardware attached to the violin, the highly customized bow imposed its use. In [7], Rasamimanana performs wireless measurements of acceleration of the bow by means of accelerometers attached to the bow, and uses force sensitive resistors (FSRs) to obtain the strain of the bow hair as a measure of bow pressure. This system has the advantage that can be easily attached to any bow. Conversely, it needs considerable post-processing in order to obtain motion information, since it is measuring only acceleration. This was carried out afterward by Schoonderwalt [8], who combined the use of video cameras with the measurements given by the accelerometers in order to reconstruct bow velocity profiles.

When looking into the literature for similar approaches to what is presented in this paper, we only find the work by Goudeseune [3], who used a commercial EMF device (same principles as the one we use in our application) for tracking some low-level movement parameters and use them for controlling some synthesis features in a performance scenario. The procedure of extracting movement or gestural parameters was not much elaborated, as he just used speeds or positions/rotations of the sensors in the violin or bow without trying to extract relevant instrumental gesture parameters.

### 3. APPLICATION CONTEXT

The main objective of the work presented in this paper is to propose a new methodology for the acquisition of violin instrumental gesture parameters from performance recordings, and also for the automatic construction of an annotated database including instrumental gesture parameters along with captured audio and video streams of the recorded performance. The analysis of the gesture data present in our annotated database will allow us to analyze instrumental gesture data in order to use them to control sample-based and physical model-based violin synthesis algorithms. Thus, we can define three main con-

texts for the work presented here, considering the first one to be the most important (it is the focus of the paper) since it is serving as a departing point for further research:

- **Database construction.** Acquisition, pre-processing, and annotation of violin performance database, including relevant instrumental gesture parameters.
- **Retrieval.** Retrieving samples from a violin performance database, based on results and annotations of instrumental gesture parameters analysis, and use selected samples in a concatenative sample-based violin synthesizer.
- **Mapping.** Controlling a violin synthesizer based on physical models, from the streams corresponding to instrumental gestures (e.g. bow speed) that have been acquired, stored and/or (re-)rendered.

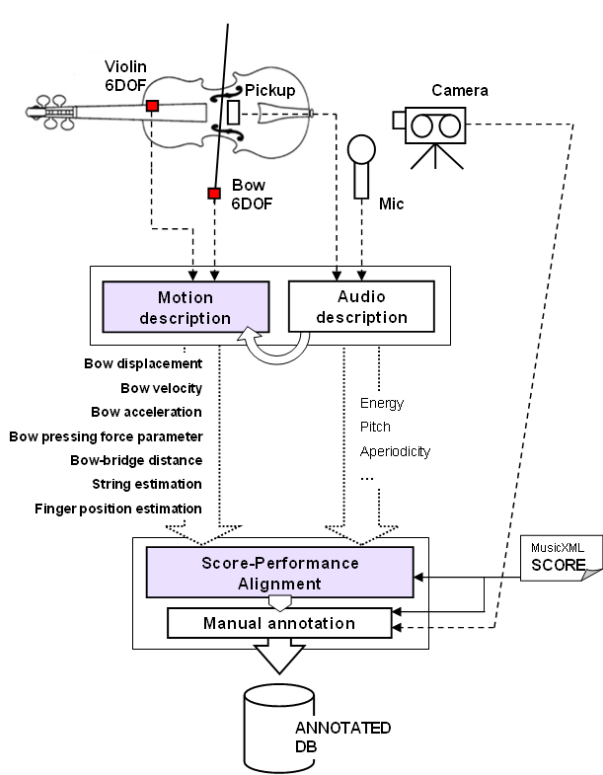
#### 3.1. Application requirements

We have to consider several requirements in our application context. First of all, since our aim is to capture the most relevant aspects of violin performance focusing on instrumental gestures, we have to define the list of parameters to measure. Extensive and detailed studies [1] [2] have shown the most important timbre related bowing parameters to be bow transversal position, bow transversal velocity, bow to bridge distance, and bow pressing force.

In addition, we consider the actual violin and bow movements to be relevant, because they also belong to gestures occurring during the performance process. Another requirement is the possibility to use the acquired data for characterizing and modelling such performance gestures, particularly instrumental gestures. Hence, the accuracy and robustness of our measurements should be as good as possible, with a high sampling rate. Furthermore, the acquisition system/procedure must prove to be non-intrusive, so that the performance process remains unaffected, allowing the performer to play as usual, and avoiding any sound modifications. Not less important is to be flexible and portable, so that it can be easily adapted to any violin with any bow, in order to capture the performer's instrumental gestures while using his own instrument. Moreover, we aim at obtaining as many instrumental gesture parameters as possible, while keeping a computationally cheap data processing.

After considering all the previous requirements and concerns, we decided to use the Polhemus Liberty<sup>1</sup> system, a 6DOF tracking system based on electro-magnetic fields (EMF), and consisting of two wired small receiving sensors ( $\phi$  0.5cm) and a transmitting source. Each sensor provides its 3DOF for position and 3DOF for orientation/rotation at 240Hz sampling rate, with static accuracies of 0.75mm and 0.15° RMS within a range of 1.5m to the source, respectively. We attach one sensor to the violin body and the other to the bow (see Figure 2) and, from their position and orientation, compute the position of the ends of

<sup>1</sup> [www.polhemus.com](http://www.polhemus.com)



**Figure 1.** Schematic representation of the acquisition process. The two main steps that are described in this paper have been highlighted: *motion description* and *score-performance alignment*.

each string and the hair ribbon at any time. From such positions, we are able to obtain a number of relevant instrumental parameters (see Section 5) extending the four stated above. In order to minimize the intrusiveness of our system, we place the bow sensor at the bow’s gravity center (see Figure 2) and attach its wire to the performer’s right forearm, so that it does not affect the performer, especially when considering its small weight of 6 gm compared to a typical bow weight of 70 gm.

### 3.2. Acquisition process overview

The acquisition process is developed in three main steps: data capturing, data processing, and database annotation. We have schematically represented acquisition data flow in Figure 1. In our performance recordings, after calibrating the measurement system (see Section 4), we obtain data from two audio sources (microphone and electric violin pickup), a video-camera, and two 6DOF sensors of the EMF tracker.

Once the raw data streams are captured and synchronized, some data processing is applied in order to obtain, in real-time, a number of relevant instrumental gesture parameters along with some audio descriptors. Finally, the annotation of the database is carried out off-line by means of score-performance alignment plus some additional manual annotation (e.g. finger being used). In this work, we put emphasis (see Figure 1) into *motion descrip-*

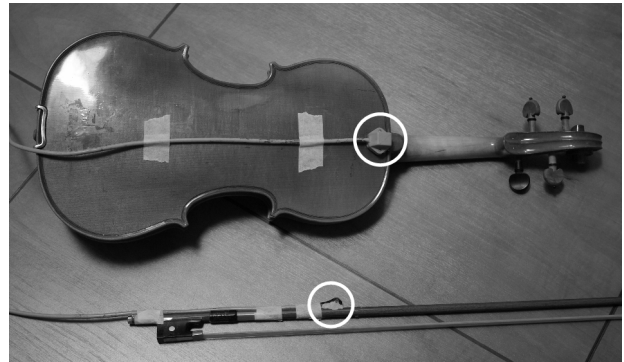
*tion* (see Section 5) and *score-performance alignment* (see Section 6).

Once the data acquisition is carried out, we store instrumental gesture data at different processing levels by following the principles and namespace extensions of the recently proposed Gesture Description Interchange Format (GDIF), as we are describing in [5].

## 4. SYSTEM CALIBRATION

The calibration process consists of three steps. The first two steps provide the means for obtaining the positions of each one of the string ends for any position/rotation of the violin. The third step will allow tracking the hair ribbon ends for any position/orientation of the bow. The whole calibration process takes no longer than 10 minutes, including the attachment of sensors to the instrument.

We are using two sensors  $S_1$  and  $S_2$ . For the whole calibration process, as well as during performance, the sensor  $S_1$  will be attached to the violin body. With respect to the sensor  $S_2$ , it will be used during the second calibration step as a marker for annotating the positions of the string ends referred to the position and rotation of the sensor  $S_1$  (placed in the violin body), while during the third calibration step, it will be placed in the bow, and will remain fixed there during performance. Sensor placements (see Figure 2) have been chosen to minimize the effect to the performance, as confirmed by several performers.

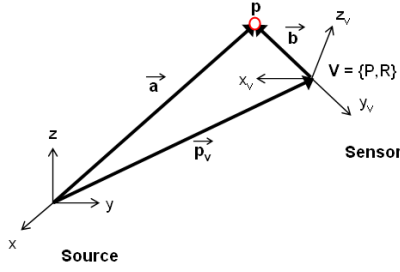


**Figure 2.** Detail of violin and bow placement of sensors during a performance recording.

### 4.1. The $\beta$ coordinates

The whole process of tracking the position of the strings and the hair ribbon is based on basic 3D algebra (e.g. translation and rotation operations). Basically, we apply translation and rotation to some known  $\beta$  coordinates (annotated during calibration) representing, for instance, string ends or hair ribbon ends positions expressed in the coordinate system of violin or bow sensors respectively. Such operations allow tracking the position of string ends or hair ribbon ends when violin or bow sensors are moving (as in performance). Next we outline the obtention of such

$\beta$  coordinates, as well as the operations used for tracking string and hair ribbon ends.



**Figure 3.** Schematic representation of coordinate systems of source and sensor, along with relevant parameters involved in the obtention of the  $\beta$  coordinates.

By looking at Figure 3, we can see two coordinate systems, corresponding to the origin and to the sensor. We can obtain the  $\beta$  coordinates in two steps. The first step is to obtain the vector  $\vec{b}$ , going from the origin of the coordinate system  $V$  to the position of the point  $p$ . For doing so, we subtract vector  $\vec{p}_v$ , which correspond to the origin of vector  $\vec{b}$ , to vector  $\vec{a}$ , as expressed in

$$\vec{b} = \vec{a} - \vec{p}_v \quad (1)$$

The second step is to rotate the vector  $\vec{b}$  so that it finally becomes expressed in terms of the coordinate system  $V$ , obtaining then the  $\beta$  coordinates. This is expressed in

$$\beta_p = M_R^{-1} * \vec{b} \quad (2)$$

where  $M_R^{-1}$  corresponds to the inverse of the rotation matrix given of  $V$  computed from the Euler angles of the rotation of  $V$ .

Now, for any new position  $P'$  and rotation  $R'$  of the coordinate system of the sensor, we are able to obtain the new position of the point  $p$ , referred to the origin coordinate system, by means of applying the new rotation matrix and translation, as expressed in

$$p' = M'_R * \beta_p + P' \quad (3)$$

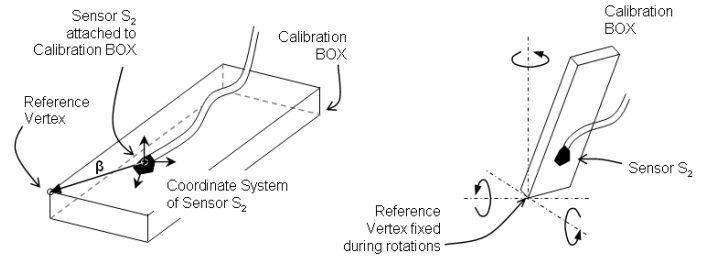
where  $M'_R$  is the new rotation matrix computed from the new Euler angles of sensor, and  $P'$  is the new position of the sensor.

#### 4.2. Sensor calibration step

This is first step of the calibration procedure. Due to the amorphous shape of the sensors, the manufacturer is not able to give details about the exact point of the sensor body for which the device is tracking position and orientation. This means a source of error at the range of the diameter of the sensor body (around  $\phi$  0.5cm). This is particularly problematic when using the sensor as a marker. Thus, we

have devised the first calibration step in order to solve this uncertainty.

We attach the sensor  $S_2$  on top of a square plastic box (we call it *calibration box*), as it is illustrated in Figure 4, and define a *reference vertex* that will be used (instead of the amorphous sensor) as a marker for annotating positions of string ends in the strings calibration step (see Section 4.3). The idea is to find the coordinates of the reference vertex in the coordinate system of sensor  $S_2$  (i.e. the  $\beta$  coordinates of the vertex). For doing so, we put the reference vertex into a small hole ( $\phi < 0.5\text{mm}$ ) on a surface (see Figure 4), and capture position and orientation data of the sensor while performing 3-axis rotations of the calibration box during some time (around 10 seconds is enough). As the distance between sensor and reference vertex remains constant, the measured positions during the capture are defining a sphere with the reference vertex as its center. Thus, we are able to find such center by applying least-squares fitting techniques<sup>2</sup>. After that, the estimated center is used for computing the  $\beta$  coordinates for every recorded frame's position  $P_{S_2}$  and rotation  $R_{S_2}$  of sensor  $S_2$ , defining (see Figure 3) point  $p$  as the estimated center, vector  $\vec{p}_v$  as going from the source to the sensor, and vector  $\vec{a}$  as going from the source to the estimated center. Then, we compute the vertex's  $\beta$  coordinates as an average.



**Figure 4.** Schematic view of the first calibration step. Sensor  $s_2$  is attached to the calibration box. The aim is to find the  $\beta$  coordinates of the reference vertex.

Now we are able to express the exact positioning of the calibration box's reference vertex in the source coordinate system (by means of eq (3)), and use it as a marker in the strings calibration step (see below).

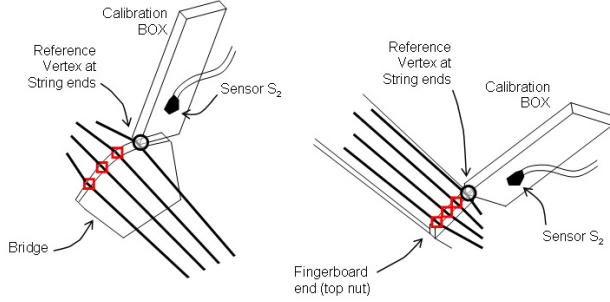
#### 4.3. Strings calibration

The string calibration step consists on the obtention of the  $\beta$  coordinates of the eight string ends, referred to the sensor  $S_1$ , placed in the violin body. For doing so, we use the reference vertex of the calibration box (see Section 4.2) as a marker. For each string end (four corresponding to the bridge ends, and another four corresponding to the finger-

<sup>2</sup> <http://www.eurometros.org/>



board ends), we annotate both position and rotation of the violin's sensor  $S_1$ , together with the position of the reference vertex. With these annotated values, we compute now the eight  $\beta$  coordinates as outlined in Section 4.1, defining (see Figure 3) point  $p$  as the end of the string, vector  $\vec{p}_v$  as going from the source to the violin sensor (sensor  $S_1$ ), and vector  $\vec{a}$  as going from the source to the reference vertex of the calibration box.



**Figure 5.** Schematic view of the first calibration step. Sensor  $S_1$  remains attached to the violin body.

Once we have obtained the  $\beta$  coordinates of each of the eighth string ends, we easily get the new position  $p'$  of each string end for a given new violin coordinate system by means of applying equation (3), where  $M'_R$  is the rotation matrix computed from the Euler angles of the sensor  $S_1$ , and  $P'$  is the position of the sensor  $S_1$ .

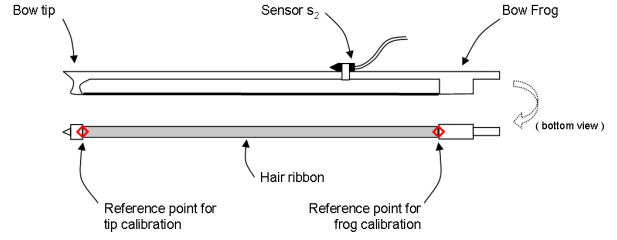
#### 4.4. Hair ribbon calibration

In the hair ribbon calibration, we first remove the sensor  $S_2$  from the calibration box, and attach it to the bow, close to the bow's center of gravity, as it is illustrated in Figures 6 and 2. Sensor  $S_2$  will remain there during performance. As it was carried out for the case of the strings, we need to obtain the  $\beta$  coordinates (now referred to the bow's coordinate system) of the hair ribbon ends. This is carried out separately in two analogous steps.

First, we position the bow in such way that the tip end of the hair ribbon gets placed at the bridge end of any string, and annotate position and rotation of both sensors  $S_1$  and  $S_2$ . With position and rotation of sensor  $S_1$ , we are able to obtain (see previous section) the position of the string end. Then, we obtain the  $\beta$  coordinate of the tip end by means of steps in Section 4.1, defining (see Figure 3) point  $p$  as the end of the string, vector  $\vec{p}_v$  as going from the source to the bow sensor (sensor  $S_2$ ), and vector  $\vec{a}$  as going from the source to string end.

Then, we repeat the process with the frog end of the hair ribbon (see Figure 6), and from now on, from the computed  $\beta$  coordinates, we are able to obtain, again by applying equation (3), the position  $p'$  of the hair ribbon ends for a given position  $P'$  and rotation matrix  $M'_R$  of the sensor  $S_2$ . We are aware of the non-zero width of the hair ribbon,

and we decided to use the middle point of the hair ribbon end (see Figure 6) during calibration.



**Figure 6.** Schematic view of the second calibration step. Sensor  $S_1$  remains attached to the violin body. We illustrate here the placement of sensor  $S_2$  on top of the bow, as well as the exact places of contact between hair ribbon ends (middle point for hair ribbon width) and the bridge end of the string.

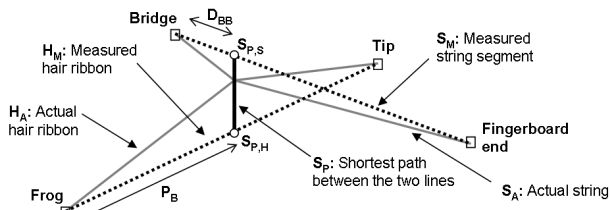
## 5. MOTION DESCRIPTION

By applying the procedure outlined in previous section, we are able to obtain both the ends of each one of the strings and the ends of the hair ribbon at any time of the performance. Now we give details on the computation methods we apply for the accurate obtention of several instrumental gesture parameters, as it is the case of a number of bowing parameters, the automatic estimation of the string being played, and others as an estimation of the finger position, or even a flag telling us whether the performer is actually playing.

### 5.1. Bowing parameters computation

Responding to our requirement to obtain relevant timbre-related bowing parameters, we compute bow transversal position  $P_B$ , bow transversal velocity  $V_B$ , bow acceleration  $A_B$ , bow to bridge distance  $D_{BB}$ , and a bow pressing force parameter  $P_F$ . All of these parameters are computed using the position of the hair ribbon ends, and the position of the ends of the string being played. These two pair of points represent two (ideally intersecting) lines in 3D space, so we base our computations on obtaining the intersection of the hair ribbon and the string being played. Assuming the string being played as a known parameter (it is estimated automatically, see Section 5.2), we first find the shortest path  $S_P$  between the ideal (no deformation, see Figure 7) hair ribbon, and the ideal (no deformation) string. Since both the string and the hair ribbon suffer deformations, the intersecting segment  $S_P$  will present a non-zero length. We find such segment by considering it to be perpendicular to both the ideal hair ribbon and the ideal string.

Once we have found the intersections of  $S_P$  with the measured string  $S_{P,H}$  and hair ribbon  $S_{P,S}$ , we are able to compute the bowing parameters as detailed next:



**Figure 7.** Measured string and hair ribbon segments, computed from their extracted end points, versus their actual configuration. Deformations have been exaggerated in order to illustrate the importance of segment  $S_P$ .

- The **bow transversal position**  $P_B$  is computed as the euclidean distance between  $S_{P,H}$  and the measured frog end of the hair ribbon.
- The **bow transversal velocity**  $V_B$  is computed as the derivative of the bow position/displacement  $dP_B/dt$ .
- The **bow transversal acceleration**  $A_B$  is computed as the derivative of the bow transversal velocity  $dV_B/dt$ .
- The **bow to bridge distance**  $D_{BB}$  is computed as the euclidean distance between  $S_{P,S}$  and the measured frog end of the hair ribbon.
- The **bow pressing force parameter**  $P_F$ , we use the modulus of the intersecting line segment  $S_P$ , which is perpendicular to the string and to the hair ribbon. This segment will get longer for higher deformations due to pressing force. Thus, we compute the euclidean distance between  $S_{P,S}$  and  $S_{P,H}$ , expressing it cm.

We are aware of the non-linear correspondence of the segment  $S_P$  length with the actual force applied by the hair ribbon to the string, although we have not yet carried out exhaustive tests in order to find some linearization function (highly dependent on bow transversal position and tilt), but rather a qualitative evaluation. For doing so, we implemented a pressing force measurement method similar to the one used in [7], and we have observed a high correlation with our pressing force parameter  $P_F$ . Therefore, we believe it might become an interesting timbre-related parameter.

## 5.2. String automatic extraction

In order to automatically detect the string being played, we have devised a method based on measuring the angle of the line joining both hair ribbon ends, and the violin plane. The procedure is outlined as follows. First, we obtain the plain of the violin from the ends of the outer strings (four points in 3D space) by means of least squares fitting techniques<sup>3</sup>. Then, we project the hair ribbon on

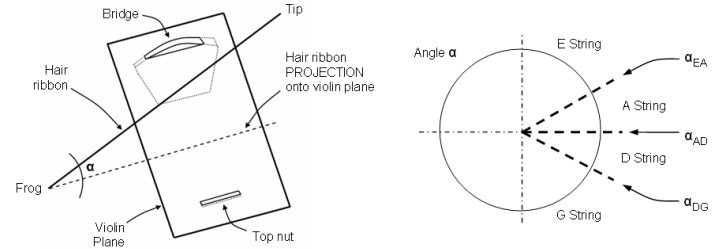
<sup>3</sup> <http://www.eurometros.org/>

the violin plane, and compute the angle between the hair ribbon and its projection, as illustrated in Figure 8. From the computed angle, we are able to obtain a good estimation of the string being played by defining sectors in a circumference, as it is also depicted in Figure 8.

For the annotation of the angle limits  $\alpha_{EA}$   $\alpha_{AD}$   $\alpha_{EG}$ , we carry out a short recording (once the system has been calibrated, as a final calibration step) asking the performer to play a known sequence of strings while we compute the angle  $\alpha$ . The sequence of strings is the following:

$$E \rightarrow E + A \rightarrow A \rightarrow A + D \rightarrow D \rightarrow D + G \rightarrow G$$

Then, we manually segment the recorded sequence and extract the limits by averaging the obtained angles along segments where two strings were played, extracting angle limits in a semi-automatic manner. For one of the recordings carried out, we obtained the followed angle limits:  $\alpha_{EA} = 15^\circ$ ,  $\alpha_{AD} = -2^\circ$ , and  $\alpha_{EG} = -19^\circ$ . Due to the asymmetric shape of both the bridge and the top nut, the angle limits are not centered around zero, both slightly shifted. In order to avoid glitches in the estimation of the string, we applied a  $2^\circ$  hysteresis cycle at the decision boundaries, resulting into a more clean signal.



**Figure 8.** Automatic detection of the string being played. Left figure illustrates the angle  $\alpha$  between violin plane and hair ribbon. Right figure depicts circumference sectors and angle limits assigned each of the strings.

Once we automatically extract the string being played, we use it for the computation of several bowing parameters (see Section 5.1). For the particular cases of bow transversal velocity and acceleration, we instead use the ends of an hypothetical fifth string placed in the middle, between string A and string D (we compute its ends from the ends of strings A and D). The reason for doing so resides on considering the computation of bow speed/acceleration to be independent of the string being played, and also on the avoidance of glitches in velocity and acceleration contours due to sudden string changes.

## 5.3. Other parameters

We obtain two additional instrumental gesture parameters from captured data. The first one is an estimation of the *finger position* by using the estimation of string being played together with the fundamental frequency contour

extracted from audio. This extracted parameter results to highly useful for correcting the fingerboard end position of the string (replacing it by the computed finger position), and therefore leading to a more accurate computation of the bowing parameters. However, we are not yet applying it for such purpose. By relating string length to fundamental frequency of string vibration, we compute the finger position as the distance  $D_{NF}$  from the top nut by means of equation (4), where  $L_S$  is the total string length,  $f_{0_S}$  is the fundamental frequency of the open string, and  $f_0$  is the fundamental frequency extracted from audio.

$$D_{NF} = L_S \left(1 - \frac{f_{0_S}}{f_0}\right) \quad (4)$$

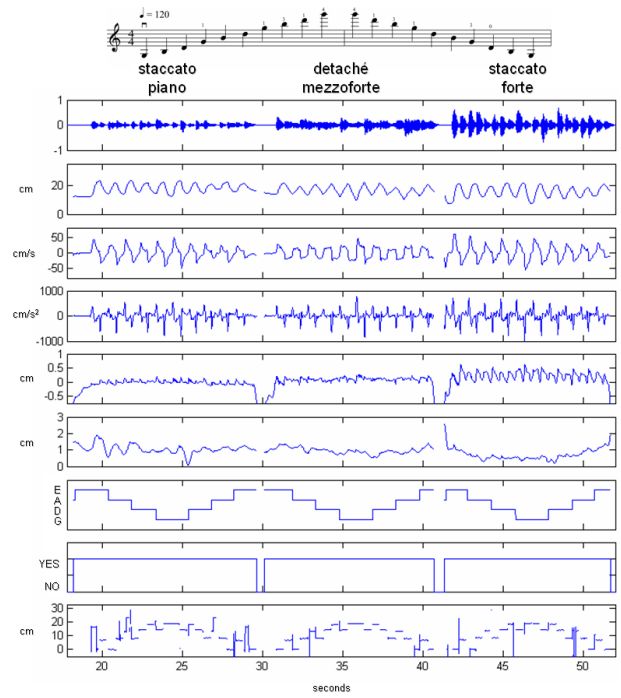
As a second additional instrumental gesture parameter, we use the distance between hair ribbon and string segments (i.e. the modulus of the segment  $S_P$ , see Figure 7 and the bow transversal velocity  $V_B$  (see Section 5.1), for computing a *playing state estimator* (playing/not playing), being helpful as a flag for recording segmentation purposes. We have decided to activate such flag when  $V_B = 0$  and  $|S_P| > 0.25\text{cm}$ , obtaining satisfactory results by applying some hysteresis at the decision boundaries.

#### 5.4. Results

It turns out to be difficult to formally assess the reliability of the whole acquisition methodology, but we give here some figures. In order to quantitatively evaluate the accuracy of the measurements, we have carried out a test in which we compare, in static conditions, real and extracted values of two bow transversal position  $P_B$  and bow to bridge distance  $D_{BB}$ . For doing so, we painted several ticks on the strings (starting from the bridge end, each cm up to 5 cm), and on the hair ribbon (starting from the frog end, each 15 cm up to 60 cm), and computed errors for several combinations bow transversal position and bow to bridge distance. We obtained average absolute errors in the range of 0.20 cm for bow transversal position, and in the range of 0.25 cm for bow to bridge distance, which we consider low enough for our research purposes. Apart from the intrinsic accuracy of the tracker (see Section 3.1), we believe that error propagation during the computations carried out in motion description is the main source of error. Figure 9 shows acquired instrumental gesture parameters for a scale played at different dynamics and articulations.

### 6. SCORE-PERFORMANCE ALIGNMENT

As a first step for the annotation of our database, we convert the performed scores (Sibelius format<sup>4</sup>) into MusicXML files including all the annotations. Then, we parse the MusicXML file and obtain all score annotations, including MIDI pitch, onset, durations, dynamics, bow direction, accents and articulation type (legato or staccato) as annotation files, by means of a custom parser that has

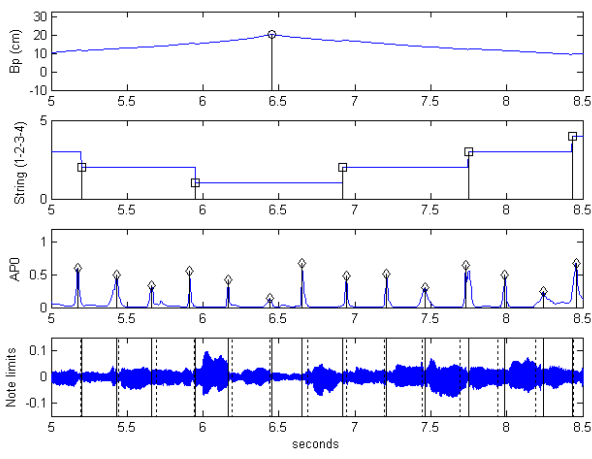


**Figure 9.** Acquired instrumental parameters for a scale played at different dynamics and articulations. From top: audio, bow position  $P_B$ , bow velocity  $V_B$ , bow acceleration  $A_B$ , pressing force parameter  $P_F$ , bow-bridge distance  $D_{BB}$ , string being played  $ST_{est}$ , *playing* descriptor, and finger position  $D_{NF}$ .

been developed for our application. Such transcription provides the nominal onset/offset times of the notes present in the score, but due to the possible deviations introduced by the performer, we need to make sure that our performance database annotations are well aligned to the score. For doing so, we correct note change times (offset/onset time of consecutive notes) coming from the score transcription, by using some of the instrumental gesture extracted parameters along with a measure of aperiodicity obtained by means of the YIN [4] fundamental frequency estimation algorithm.

The procedure can be summarized as follows. First, we define a window  $\omega_{tc_i} = \{tc_i - \Delta_{l,i}, tc_i + \Delta_{r,i}\}$  around each note change time  $tc_i$  found in the score transcription. Assuming that the performer deviations on note change times will not exceed half the duration of the pair of notes involved (preceeding and subsequent), we define the left and right widths  $\Delta_{l,i}$  and  $\Delta_{r,i}$  of the time window as being half of the duration of the preceeding and subsequent notes respectively. Within the time window  $\omega_{tc_i}$ , we find the times of any bow direction change  $t_{b_c}$  from zero-crossing of bow velocity  $V_B$  (see equation (5)), any string change  $t_{st_c}$  from local maximum of absolute value of derivative of the string estimation  $ST_{est}$  (see equation (6)), and pitch transitions  $t_{f_0_c}$  from local maxima a measure of the aperiodicity function  $AP_0$  obtained by means of the YIN [4] fundamental frequency estimation algo-

<sup>4</sup> www.sibelius.com



**Figure 10.** Detected bow direction changes, string changes, and pitch transition changes. From top: bow position  $B_P$ , string estimation  $ST_{est}$ , and aperiodicity function  $AP_0$ . At the bottom we observe the audio signal with note transitions, dashed for nominal times and solid for detected times.

rhythm (see equation (7)). We have illustrated in Figure 10 the detections obtained from a recording excerpt.

$$t_{bc} = \arg_t V_B(t) = 0 \quad (5)$$

$$t_{stc} = \arg_t \left| \frac{\delta}{\delta t} ST_{est}(t) \right| \neq 0 \quad (6)$$

$$t_{f0c} = \argmax_t AP_0(t) \quad (7)$$

Once the detection process has been carried out, we annotate the obtained change time for each time window. For those windows in which several of the three indicators of note change led to a note change detection, we give priority to the bow direction changes, followed by string changes, and finally by fundamental frequency transitions. We have not yet carried out a formal evaluation of the proposed alignment methodology, but we have revised the obtained corrections, observing satisfactory improvements to the score transcription note boundaries.

## 7. CONCLUSION

We have presented a new method for real-time acquisition of violin instrumental gesture parameters using a commercial two-sensor 3D tracking system based on electromagnetic field (EMF) sensing. We have paid attention to intrusion, flexibility, and robustness. Our approach, able to obtain a number of relevant instrumental gesture parameters with low-cost processing of the raw data coming from the tracking system, has proved to be little intrusive for the performer, while being easily ported to any violin or bow, allowing each performer to use her/his own instrument. We have detailed calibration, instrumental gesture parameters computation, and the score-performance

alignment procedure used for annotating our performance database.

The methodology outlined here serves us for the automatic creation of a violin traditional performance database. We are already developing methods for applying the obtained results (annotated database) to high-level analysis and modeling of violin instrumental gestures. Developed models will allow us in the near future to drive a sample-based concatenative violin synthesizer, while rendering of instrumental gesture parameter contours will enable to control physical model-based violin synthesis. We also aim at designing a system capable of automatically recognizing violin instrumental gestures from acquired parameters. As a main improvement of the methodology, we plan to carry out a formal study of the correlation between our pressing force parameter with the actual force applied to the bow.

## 8. ACKNOWLEDGEMENTS

This work has been supported by Yamaha Corp. We thank violinist Oriol Saña for his long performances. We would also like to thank Carlos Spa and Nils Peters for their disinterested help in calibration and accuracy tests, and to Cristina González for reviewing the manuscript.

## 9. REFERENCES

- [1] A. Askenfelt. Measurement of bow motion and bow force in violin playing. *Journal of the Acoustical Society of America*, 80, 1986.
- [2] A. Askenfelt. Measurement of the bowing parameters in violin playing. ii. bow-bridge distance, dynamic range, and limits of bow force. *Journal of the Acoustical Society of America*, 86, 1989.
- [3] C.M.A. Coudeseune. Composing with parameters for synthetic instruments. *PhD Thesis, Univ. Illinois*, 2001.
- [4] A. de Cheveigné and H. Kawahara. Yin, a fundamental frequency estimator for speech and music. *Journal of the Acoustical Society of America*, 111:4, 2002.
- [5] E. Maestre, J. Janer, A.R. Jensenius, and J. Malloch. Extending gdiff for instrumental gestures: the case of violin performance. *International Computer Music Conference*, Submitted, 2007.
- [6] J. A. Paradiso and N. A. Gershenfeld. Musical applications of electric field sensing. *Computer Music Journal*, 21, 1997.
- [7] N. Rasamimanana. Gesture analysis of bow strokes using an augmented violin. *IRCAM DEA ATIAM*, 2003.
- [8] E. Schoonderwaldt, N. Rasamimanana, and F. Bevilacqua. Combining accelerometer and video camera: Reconstruction of bow velocity profiles. *Conference on New Interfaces for Musical Expression*, 2006.
- [9] D. S. Young. The hyperbow controller: Real-time dynamics measurement of violin performance. *Conference on New Interfaces for Musical Expression*, 2002.
- [10] D. S. Young. Wireless sensor system for measurement of violin bowing parameters. *Stockholm Music Acoustics Conference*, 2003.