Informing bowing and violin learning using movement analysis and machine learning

Erica Volta, Paolo Alborno and Gualtiero Volpe

Casa Paganini – InfoMus, DIBRIS, University of Genova, Genova (Italy)
erica.volta@edu.unige.it, paoloalborno@gmail.com, gualtiero.volpe@unige.it

Abstract. Violin performance is characterized by an intimate connection between the player and her instrument that allows her a continuous control of sound through a sophisticated bowing technique. A great importance in violin pedagogy is, then, given to techniques of the right hand, responsible of most of the sound produced. This study analyses the bowing trajectory in three different classical violin exercises from audio and motion capture recordings to classify, using machine learning techniques, the different kinds of bowing techniques used. Our results show that a clustering algorithm is able to appropriately group together the different shapes produced by the bow trajectories.

Keywords: Movement Analysis, Music Performance, Music Learning, Machine Learning, K-Means, Multimodal Interactive Systems

1 Introduction

1.1 Background

One of the central elements in violin pedagogy is the bowing technique. Methods introducing violin pedagogy commonly begin illustrating how to hold properly violin and bow, as well as the proper posture to support body movements without impeding bow arm (Galamian, 1962). A great importance in the violin pedagogy is given to techniques of the right hand that is responsible of most of the sound produced.

Several studies investigated the movement of the bow, finding a strict relationship between motion characteristics and quality of the performance, e.g., the bowing motion should be fluid (Galamian, 1962) and circular (Starker, 1979).

One of the first and most influential technical approaches to the study of bow movement is Hodgson's *Motion Study and Violin Bowing,* published in 1934. In his famous work, Hodgson used early methods of photographic motion tracking to study the circular nature of bowing technique in cyclographs (see Figure 1).

The controversial insights of Hodgson's work, showing that the bow's trajectory is always curved, has caused an animated pedagogical debate, but the knowledge of the curved nature of bowing has influenced the pedagogy of the last century, giving to violinists an explanation and a metaphor to understand the correctness of their movements, since it is generally not common for students or teachers to see their own playing movements represented in a visual way.

The body of scientific research related to music education has grown significantly in recent decades, e.g., see (Rauscher, Shaw, & al, 1997), as well as the trend in developing sensors-based systems to use bow gestures in interactive performance (Machover, 1992), (Nichols, 2002), (Overholt, 2005). Despite this growth, since Hodgson's work, technology has been rarely applied to music pedagogy and usually restricted to other domains, such as to audio and video recording and playing.

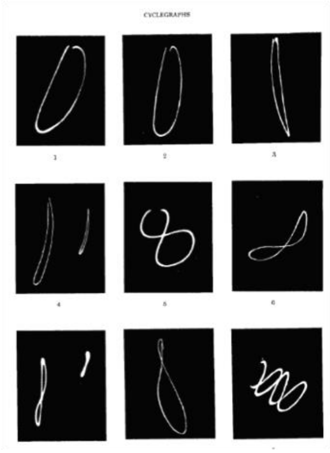

Figure 1: Hodgson's cyclographs

This study is carried out in the framework of the EU-H2020-ICT Project TELMI, having the purpose of enhancing music learning through the development of multimodal systems for real-time and off-line feedback to students.

The aim of the present preliminary work is to explore whether multimodal systems and machine learning techniques can be used for analyzing bow trajectories as a means of contributing to music performance pedagogy, by working on selected recordings of renowned performers and teachers recruited by the Royal College of Music in London.

1.2 Bowing techniques

Bow control is a central musician's skill, giving the violinist the ability to direct bow's motion during playing.

In his work, Hodgson divided bowing movements in three categories of motion:

- Movements across the string, influenced by tilt, speed and contact point techniques;
- Rotation movement around the string, that allows changing across strings and changing in direction;
- Movement towards and away from string, variating the weight and that is responsible of particular articulation effects.

For this work, we chose to focus our attention on articulation techniques, as one of the most delicate parts of violin education and as one of the elements that a multimodal system can help to analyze.

From the entire TELMI archive (see Section 1.3) we selected the *Martelé* (from Kreutzer, Op.7), *Spiccato*, and *Sautillé* (from Ševčík, Op.3) exercises recorded by four internationally renowned and esteemed professional violinists involved in the TELMI project. The choice of these exercises was made by considering the importance of these three different bowing techniques and the differences between them that are often confused and difficulty master by students.

One of the difficulties of these studies is related to speed, because there is a common ground where one should be able to make the change from *Spiccato* to *Sautillé* and vice versa without changing any character of the sound profile. Furthermore, the mechanic of these two bowings are completely different. In *Spiccato* every single note is played actively, whereas in *Sautillé* the jumping activity is left quite exclusively to the resiliency of the stick. A further difference lies in the hand's motion: according to Hodgson, in fact, during *Spiccato* the bow designs an eight in the air and when *Spiccato* is quickened to *Sautillé*, the movement of the hand changes to an ellipse, but the bow continues to draw an eight in the air. *Martelé* represents, finally, a third type of fundamental movement of bowing, since it is at the basis of essential bowing techniques, such as *Staccato*, where the pressure is released between each stroke, and the bow speed has to be quite fast, yet light.

1.3 The TELMI archive of multimodal recordings

One of the milestones of TELMI is to build a corpus of multimodal data for informing the development of a multimodal interactive system for technology-enhanced violin learning and teaching (Volpe, Kolykhalova, Volta, & al., 2017).

The archive is organized in a structured collection of exercises that follow the learning path of classical violin students.

It includes several sources of data, such as motion capture of the performer, of the violin and of the bow, ambient and instrument audio, video, physiological data, (electromyography) and Kinect data.

The corpus of material consists of 41 exercises, concerning handling the instrument, techniques of the right and left hands, articulation, and some expressive works (such as Elgar, *Salut d'amour, Op. 12*).

All the recorded data were synchronized and played back using the EyesWeb XMI[1] platform (Camurri, Hashimoto, Ricchetti, & al., 2000), (Volpe, Alborno, & al., 2016).

2 Recordings and segmentation

We recorded 4 players performing the three selected exercises (*Martelé*, *Spiccato* and *Sautillé*) during the recording sessions for the TELMI archive. The violinists received the entire list of exercises in advance and the use of music sheets was allowed.

A Qualysis motion capture system endowed with 13 video cameras was used to record each performance. MoCap data was recorded at 100 Hz and synchronized with two video streams at 50 fps and with a pickup microphone stuck on the violin. The bow was endowed with 4 lightweight reflective markers. 2 more markers defined as virtual markers, together with 6 rigid bodies were included to enhance tracking robustness and reliability. On the violin, we attached 6 further markers.

After the recordings, data was segmented, by considering the musical structure, to isolate single bowing movements for each music phrase. We extracted 869 segments in total from 12 recordings we considered.

Using EyesWeb we computed the 3D-trajectories of the tip of the bow, to check the presence of Hodgson's bowing shapes.

[1] www.casapaganini.org/eyesweb

We then computed six different features on each obtained segment, in particular:

- *Acceleration*
- *Trajectory length*
- *Kinetic energy*: an approximation of the overall energy spent while performing a movement with the bow. It is computed as the total amount of displacement in all of the tracked points.
- *Curvature*: The derivative of the position vector over the curve of the trajectory provides a *tangent vector* to the curve. The curvature is the derivative of such a vector, and describes how the tangent vector changes. As an example, a trajectory following the contour of a geometric shape, such as a square, will bend sharply in some points, so its curvature will have high values.
- *Directness*: this is a measure of the extent to which a given trajectory is direct or flexible. It is computed as the ratio between the Euclidean distance calculated between the starting and the ending point of the considered trajectory, and its length.
- *Smoothness*: this corresponds to the third derivative of the position and it has often been used as a descriptor to evaluate how a motion trajectory varies "slowly" over time (Flash and Hogan, 1985).

3 Clustering

To obtain the same number of features for each segment, we extracted a cumulative histogram with 25 bins for each considered feature, resulting into a 150-dimension feature vector data.

K-means was applied to all the feature vectors, to figure out whether the different kinds of articulation exercises we were studying can be distinguished from the characteristics of the bow motion.

We then estimated the best number of clusters (i.e., the k parameter to seed the k-means algorithm), that is shown in the Elbow curve in Figure 2.

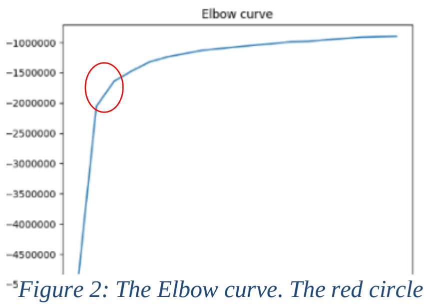

Figure 2: The Elbow curve. The red circle identifies the number of clusters (k =3) we choose for this study.

A value of k=3 was detected as the most appropriate value. Finally, we applied a PCA reduction to lower the dimension to the most representative 2 and 3 features, and visualized the clusters. Resulting data clusters are shown in Figure 3 with different colors. As the figure shows, the considered segments were split mainly in three clusters. We verified a clear separation between three particular bow motion trajectories, i.e., segments belonging to the same cluster present similar trajectories.

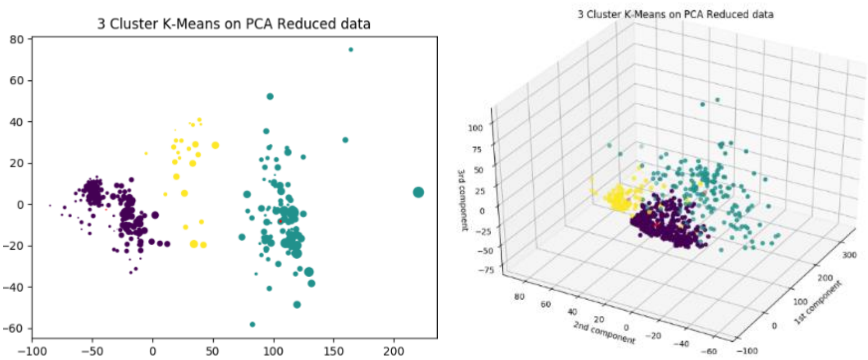

Figure 3: 3 Clusters obtained by applying K-Means on PCA reduced data

We observed that most of the segments composing the purple cluster have been extracted from *Sautillé* performances, and segments belonging to the yellow cluster are mostly related to the *Spiccato* pieces. The most spread cluster, the light green one, is mostly composed by segments belonging to *Martelé* recordings.
Our hypothesis, that needs to be investigated in the future with an extension of the presented work, is that the set of features we computed on the trajectories can effectively be used to distinguish different bowing exercises and articulation techniques.

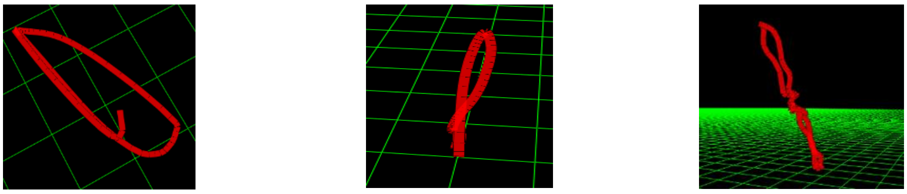

Figure 4: 3D visualization of the trajectories "drawn" by the bow while performing three different exercises. From left to right: a) Martelè: characterized by a circular trajectory, b) Sautillè: characterized by an "8"- trajectory, c) Spiccato: characterized by a lace trajectory.

4 Conclusions

In this paper, we presented some preliminary results of our analysis of the TELMI multimodal corpus of data. Results need to be confirmed by a deeper analysis of the clustering we obtained. We aim at developing more sophisticated techniques to realize an adaptive system, able to understand the type of bow movement and violin exercise starting from movement features.

The use of machine learning techniques in a music educational project aims to further develop algorithms able to automatically assess the quality and precision of the music performance to help students to enhance their musical skills.

5 Acknowledgments

We thank violinists Madeleine Mitchell, John H. Gilbert, Eulalie Charland, and Berent Korfker for participating in the recordings. We thank luthier Alberto Giordano for providing and preparing the violin and the bow.

This work is partially supported by the EU-H2020-ICT Project TELMI.

TELMI has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 688269.

References

Brandfonbrener, A. (2004). Healthier music students: Can medicine and music prescrobe in concert? *Medical Problems of Performing Artists*, 19(1). 1-2.

Camurri, A., Hashimoto, S., Ricchetti, M., et al., (2000). EyesWeb: Toward Gesture and Affect Recognition in Interactive Dance and Music Systems. *Computer Music Journal*, 57-69.

Galamian, I. (1962). *Principles of Violin Playing and Teaching.* London: Faber & Faber.

Kolykhalova, K., Volta, E., Ghisio, S., et al., (2017). A multimodal corpus for technology-enhanced learning of violin playing. *Proceedings of the 12th biannual Conference of the Italian ACM SIGCHI Chapter.* CHItaly.

Machover, T. (1992). *Hyperinstruments – a progress report 1987-1991.* Technical report, MIT Media Laboratory.

Nichols, C. (2002). The vBow: A virtual violin bow controller for mapping gesture to synthesis with haptic feedback. *Organized Sound*, 7(2):215-220.

Overholt, D. (2005). The overtone violin. *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression (NIME05)*, (pp. 34-37). Vancouver, BC, Canada.

Rauscher, F. H., Shaw, G. L., et al, (1997). Music causes long-term enhancement of preschool children's spatial-temporal reasoning. *Neurological Reserch*, 19, 2-8.

Starker, J. (1979). An organized method of string playing. In M. A. Grodner, *Concepts in String Playing: Reflections by Artisti-Teachers at the Indiana University School of Music* (pp. 133-55). Bloomington: Indiana University Press.

Volpe, G., Alborno, P., et al., (2016). Designing Multimodal Interactive Systems using EyesWeb XMI. *Proceedings of the First International Workshop on Smart Ecosystems AVI 2016*, (pp. 49-56). Bari.