Automatically measuring biomechanical skills of violin performance: an exploratory study

Erica Volta
University of Genoa
Genoa, Italy
erica.volta@edu.unige.it

Maurizio Mancini
University of Genoa
Genoa, Italy
maurizio.mancini@unige.it

Giovanna Varni
LTCI, Télécom ParisTech, Université Paris Saclay
Paris, France
giovanna.varni@telecom-paristech.fr

Gualtiero Volpe
University of Genoa
Genoa, Italy
gualtiero.volpe@unige.it

ABSTRACT

This evaluation study explores how automated movement analysis can be used to catch the biomechanical skills needed for a physically accurate violin performance, maximizing efficiency and minimizing injuries. Starting from a previously recorded multimodal dataset, we compute movement features from motion captured data of five violinists performing three violin exercises: octave shift, string crossing, and a Romantic repertoire piece. Three violin teachers were asked to evaluate audio, video, and both audio and video stimuli of the selected exercises. We correlated their ratings with automatically extracted movement features. Whereas these features are purely visual (i.e., they are computed from motion captured data only), we asked teachers to also evaluate audio because it can be considered as the direct translation of movement skills into another modality. In this way, we can also look at possible relations between evaluation of the audio aspects of the performance and biomechanical skills of violin playing. Results show that the proposed movement features can be partially used to measure the biomechanical skills of the violin players to support learning and mitigate the risk of injuries.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; *Empirical studies in HCI*; • **Applied computing → Performing arts**; **Education**;

KEYWORDS

Movement Analysis, Music Performance, Music Learning, Multimodal Interactive Systems, 3-D motion analysis, Biomechanical modelling, Violin pedagogy, Motor control, Motor modelling

MOCO '18, June 28–30, 2018, Genoa, Italy

ACM ISBN 978-1-4503-6504-8/18/06…$15.00
https://doi.org/10.1145/3212721.3212840

ACM Reference Format:
Erica Volta, Maurizio Mancini, Giovanna Varni, and Gualtiero Volpe. 2018. Automatically measuring biomechanical skills of violin performance: an exploratory study. In *MOCO '18: 5th International Conference on Movement and Computing, June 28–30, 2018, Genoa, Italy.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3212721.3212840

1 INTRODUCTION

In literature, the interest in music performance has grown, considering several contributions from neuroscience, e.g., [1, 8–10] to music pedagogy, e.g., [11, 16, 17] exploring music performance and learning. An emerging literature investigates how technology, and in particular full-body and motion analysis technologies, may enhance music performance and learning outcomes, by minimizing at the same time the risk of injuries [13, 14]. Playing violin is mostly based on a master-apprentice relationship and it is difficult because student's interaction with the teacher is often restricted to the time of the weekly lectures, followed by long periods of self-study. This make the teacher's feedback to be dissociated from the on-line proprioceptive and auditory sensations accompanying the performance [20]. In addition, traditional teaching methods of the biomechanics components of musical performance may be based on subjective and vague perception, rather than on accurate understanding of the principles of human movement [4].

Moreover, musical performance shares many characteristics, including health risks, in common with other skill-oriented activities [15]. It is reasonable to postulate that some methodologies successfully used in sports medicine could be useful in studying the biomechanical aspects of music performance. These understandings include the motor learning theory, learning models, and technology-based system to analyse and validate the effects of training [12]. It is, e.g., a common expectation that students engage in long hours of repetitive practice to acquire technical skills, often not considering the efficacy of such practice and its health implications due to many repetitions of the same movements. Despite many individual success stories, documented rates of injuries among musicians challenge traditional methods [7].

This study aims at exploring whether methodologies and technologies from movement analysis can be applied to investigate aspects of the biomechanics of violin performance. These kind of methodologies applied to music performance and learning is not completely new. In [21] the authors applied motion capture technologies and quantitative analysis of motion to the analysis of prototypical gestures in music conduction. We exploited automated movement analysis to capture the biomechanical skills needed for a physically accurate violin performance, aiming at identifying features able to maximize efficiency and minimize injuries. Starting from a multimodal dataset of previously recorded performances, we selected motion captured data of five violinists performing three violin exercises: octave shift, string crossing, and a Romantic repertoire piece. We asked three experts to evaluate the recorded performance for the three exercises, by providing their ratings for 2 audio (i.e., intonation and note production) and 3 movement (i.e., shoulders dynamic and position, and trunk dynamic) features. We finally explored whether and how movement features extracted from full-body motion captured data correlate with the ratings provided by experts.

2 METHOD

This study is based on two grounded concepts: motor learning and motor control. In [2], theory of motor learning identifies three elements needed to success: the characterization of the skills to be acquired; skills transfer between dissimilar systems, and skills acquisition without injuries. Characterizing skills to be acquired involves a scientific analysis and understanding to identify and describe motor control patterns, such as the coordination of neural and muscoloskeletal systems. The motor patterns of professional players can be generalized and used as references for a model, to facilitate the identification of skills to be transferred from professional to student musicians. By directing attention to specific motions patterns, learners can assimilate the new skills into their technique efficiently and effectively. In our study, we analyze the data of a multimodal archive of recordings to study movement patterns of shoulders, elbows, and hips. Looking at the variation of angles between such parts of the body, we hypothesize how muscular and posture patterns affect music performance in terms of pitch intonation and dynamic.

3 THE TELMI MULTIMODAL ARCHIVE

The TELMI Project[1] is an ongoing three years H2020-ICT European Project, aimed to design and implement a multimodal interaction system for music learning, providing the user with assistive, self-learning, augmented-feedback, and social-aware prototype interactions. Such a system is conceived to be complementary to traditional music pedagogy. The project grounds on a tightly coupled interaction between technical and pedagogical partners, to implement new multimodal interaction prototypes for music learning and training, based on state-of-art audio processing, computer vision, and motion capture technologies. To reach its objectives, the TELMI Project built a corpus of multimodal data [19], structured as a collection of exercises to follow the learning path of classical violin programmes. The corpus 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 music material encompasses 41 exercises, subdivided in handling the instrument, techniques of the right and left hands, articulation and some expressive works

[1] http://telmi.upf.edu

Figure 1: Example of the segmentation we made to obtain the evaluation stimuli

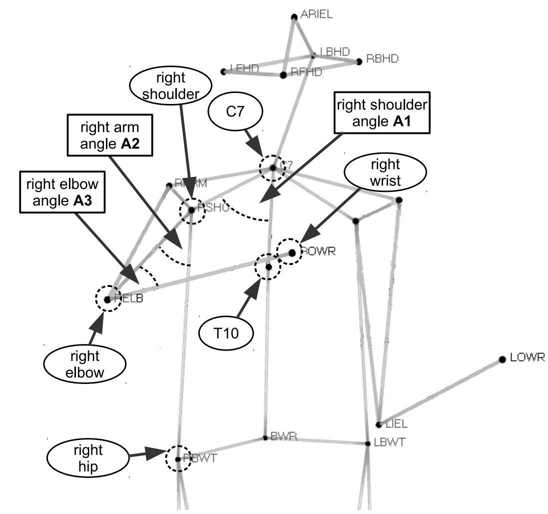

Figure 2: The joints (written in rounds) and angles (written in rectangles) we exploit for extracting movement features.

(e.g., Elgar, Salut d'amour, Op. 12). The score of an excerpt is reported in Figure 1. All the recorded data were synchronized and played back using the EyesWeb XMI platform [5], [18].

4 RECORDING AND SEGMENTATION

We focused on 5 players performing three selected exercises (Octave shift from Yost System for Violin, String Crossing from Kreutzer Etude n.13, and Salut d'amour by Edgar, see Figure 1). The violinists received the entire list of exercises in advance and the use of music sheets was allowed. The musicians were 4 professional players and one high-level non-professional violinist. After the recordings, the data was segmented, using audio and music sheet information, to isolate single music phrases. In such a way, we extracted 21 segments from the twelve recordings we considered. Using the EyesWeb XMI platform, we computed the measures of the angles between right shoulder and elbow, left shoulder and elbow, the two shoulders, and shoulder-elbow and hip.

5 MOVEMENT FEATURES

Considering the theoretical model described in [14], we focused our analysis on full-body movement features, corresponding to the ones taken into account by violin teachers to evaluate their student's performance.

Table 1: Spearman's correlations between the ratings provided by the raters and the automatically computed features. Significance level: * for $p < 0.05$, ** for $p < 0.01$

Stimulus	**Shoulders' dynamics**		**Body joints' angles**		**Trunk's dynamics**	
	rho	p-value	rho	p-value	rho	p-value
Audio	.43	.05*	.04	.87	-.03	.91
Video	.55	.008**	-.24	.28	.11	.63
Audio-Video	.61	.005**	-.07	.77	.17	.46

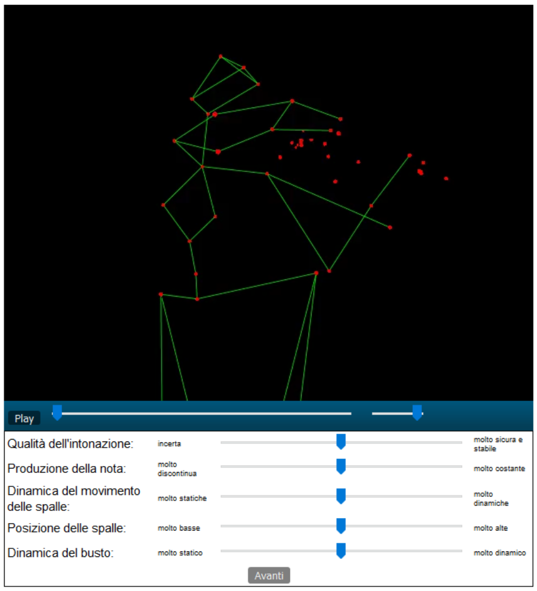

Figure 3: Web interface for evaluation of biomechanical violin playing skills: "Quality of intonation" ("Qualità dell'intonazione"), "Note production" ("Produzione della nota"), "Shoulders' dynamics" ("Dinamica del movimento delle spalle"), "Shoulders' position" ("Posizione delle spalle"), "Trunk's dynamics" ("Dinamica del busto")

According to such model, in order to achieve a better performance and minimize the rick of injuries, one could ask a performer "to consciously choreograph movements that minimize static conditions or postures". Moreover, "left elbow should have a degree of lateral movement to accommodate playing on different strings". Also, players that tend to assume unbalanced (and physically risky) postures usually "locked their shoulders in a highly abducted position". Some expert violin teachers we interviewed before computing movement features confirmed that "the fight against the tendency to raise one's shoulders is a constant in the teaching, as well as an essential prophylaxis against the onset of very boring and often very damaging professional diseases". Provided the above background experience, we identified a small initial set of movement features mainly based on (i) angles between joints and (ii) energy of joints movement. The joints and angles we looked at are highlighted in Figure 2 and include:

- joints: C7 (base of the neck), T10 (trunk vertebra), shoulder, elbow, wrist
- angles: the angle between the line C7-T10 and the line C7-shoulder (A1 in Figure 2); the angle between the line shoulder-hip and the line shoulder-elbow (A2 in Figure 2); the angle between the line shoulder-elbow and the line elbow-wrist (A3 in Figure 2).

The above joints and angles are considered both for the right and the left side of the player's body (i.e., angles A1-3 will be computed on both sides). The 3 movement features we compute on the above joints and angles are: **F1, shoulders dynamic** - we compute the Kinetic Energy of left/right shoulder and the first derivative of angles A2 and A3, and we sum up these 3 quantities; the mean between the right and the left side of the body is then calculated, obtaining a single value for F1; **F2, shoulders position** - we compute the right/left mean angle A1; **F3, upper body dynamic** - we sum up the Kinetic Energy of upper body joints C7, T10, and right/left shoulder. F1-3 are computed frame-by-frame on the 21 segments described in the previous section. In this exploratory work, we computed the average of F1-3 on each segment, obtaining a single value of each feature for that segment.

6 EVALUATION

Three expert musicians, two violin teachers with more than 20 years of experience in teaching and a professional violin performer, were recruited via email advertisements at academies of music in Italy and France. They were asked to mark their ratings on violin skills, through a questionnaire administered via an ad-hoc web interface (see Figure 3). The raters provided their answers to each question by dragging an interactive slider having a resolution of 101 values (the experts did not perceive these values, they perceived a continuous evaluation). The questionnaire consisted of 5 bipolar items ($\alpha = 0.80$) arranged in two sub-scales: a sub-scale related to the acoustic component of violin playing skills (2 items, $\alpha = 0.72$), and a sub-scale related to the movement component of violin playing skills (3 items, $\alpha = 0.70$). Internal consistency was assessed through Cronbach's α. Results show an acceptable reliability both for the whole scale and for each sub-scale [6]. The items were: pitch intonation and note production for audio, shoulders' dynamics and position and trunk's dynamics for video. 21 audio/video and audio-video stimuli for a total amount of 63 stimuli were administered in a randomized way. When the stimulus was only audio or video, raters were asked to mark their answers only on the corresponding audio or video subscale. Raters were explicitly asked to fully complete the questionnaire. The ratings given by each rater were summed up together over the items.

6.1 Analysis and results

Inter-rater reliability was assessed through a two-way random, consistency, average-measures ICC (McGraw & Wong, 1996). It enables to assess the degree at which the 3 raters provided their agreement on violin playing skills across stimuli. The results obtained for the whole scale and each sub-scale were ICC=0.81 (audio-video, 95% CI [0.60,0.92], F(19,38) = 5.24 ,$p < .001$), ICC=0.82 (video, 95% CI [0.54,0,90], F(21,24.4) = 5.68, $p < .001$), ICC= 0.72 (audio, 95% CI [0.24 -0.85], F(20,19.1) = 3.59, $p < .001$), respectively. These high values, falling in the range good and excellent (Cicchetti 1994), indicate that a minimal amount of measurement error was introduced by the 3 independent raters. The confidence interval for audio is very large and one possible explanation we found rely on the cognitive dynamic of audio perception: since the audio sensory modality, in typical people, is informed by visual cues, the absence of visual stimuli could lead to more unsure answers during the evaluation [3]. Pearson's and Spearman's correlations were run to determine the relationship between the perceived violin playing skills and the automatically computed movement features. The analysis showed that there was a strong positive correlation between the shoulders' dynamics and the perceived violin playing skills, independently of the modality of the stimuli. Spearman's correlations showed slightly higher values meaning that there is some deviation from a linear increasing monotonic relationship between the variables. Table 1 summarizes the results of the Spearman's correlations for each automated movement features.

7 DISCUSSION AND CONCLUSIONS

Results show a significant correlation between the evaluation of the experts on the violin playing skills perceived through the different modalities and the measured feature on shoulders' dynamics. This is promising in view of designing systems for music teaching and learning support. Moreover, results show that all the evaluated modalities (the visual one in particular) can provide useful feedback for both teachers and students on the biomechanical skills needed for a good violin performance. Improvements can be expected, since some of the results are not significant. For example, the measure of shoulders' position did not correlate with the perceived violin playing skills. This can be motivated by observing that all evaluators highlighted that the violinists exhibited high-locked shoulders.In the future, a similar evaluation could be repeated, using more recordings of different types and violinists to check the extent to which results generalize. Moreover, shoulders position seems to be performance-dependent, in the sense that the right shoulder rotates vertically according to the string that is currently played. Our algorithm does not take this into account, while human observers did it during the evaluation. The last feature, trunk's dynamics, does not correlate with the experts' ratings too. This result can be explained by the fact that violinists movements during the recordings were limited by the mocap suite and EMG sensors. This non ecological setup may lead to less dynamic movements. In the future, besides improving our algorithms and repeating the evaluation with violinists with more variable skills, we also plan to extend the evaluation, using a more ecological setup, based on portable and non-intrusive devices, such as range imaging sensor.

ACKNOWLEDGMENTS

We thank violinists M. Mitchell, J. H. Gilbert, E. Charland, and B. Korfker for participating in the recordings. We thank luthier A. Giordano for providing and preparing the violin and the bow. This work is partially supported by the EU-H2020-ICT Project TELMI (Grant Agreement No. 688269).

REFERENCES

[1] Ryan P. Atherton, Quin M. Chrobak, Frances H. Rauscher, Aaron T. Karst, Matt D. Hanson, Steven W. Steinert, , and Kyra L. Bowe. 2018. Shared Processing of Language and Music. *Experimental Psychology* 65 (2018), 40–48.

[2] R. Ballreich and W. Baumann. 1996. *Grundlagen der Biomechanik des Sports*. Stuttgart. (The basics of biomechanics in sports).

[3] Marilyn G. Boltz, Brittany Ebendorf, and Benjamin Field. 2009. Audiovisual interactions: The impact of visual information on music perception and memory. *Music Perception: An Interdisciplinary Journal* 27, 1 (2009), 43–59.

[4] A. G. Brandfonbrener. 2004. Healthier music students: Can medicine and music prescribe in concert? *Medical Problems of Performing Artists* 19, 1 (2004), 1–2.

[5] Antonio Camurri, Shuji Hashimoto, Matteo Ricchetti, Andrea Ricci, Kenji Suzuki, Riccardo Trocca, and Gualtiero Volpe. 2000. Eyesweb: Toward gesture and affect recognition in interactive dance and music systems. *Computer Music Journal* 24, 1 (2000), 57–69.

[6] Lee J. Cronbach. 1951. Coefficient alpha and the internal structure of tests. *psychometrika* 16, 3 (1951), 297–334.

[7] Martin Fishbein, Susan E. Middlestadt, Victor Ottati, Susan Straus, and Alan Ellis. 1988. Medical problems among ICSOM musicians: overview of a national survey. *Med Probl Perform Art* 3, 1 (1988), 1–8.

[8] Wilfried Gruhn and Frances Rauscher. 2002. The neurobiology of music cognition and learning. *The new handbook of research on music teaching and learning* (2002), 445–460. http://www.uwosh.edu/psychology/faculty-and-staff/frances-rauscher-ph.d/GruhnRauscher_2008.pdf

[9] Robert Zatorre Isabelle Peretz. 2003. *The Cognitive Neuroscience of Music*. Oxford University Press, Oxford.

[10] L. James and Jr. Reifinger. 2018. The Relationship of Pitch Sight-Singing Skills With Tonal Discrimination, Language Reading Skills, and Academic Ability in Children. *Journal of Research in Music Education* 0, 0 (2018), 0022429418756029. https://doi.org/10.1177/0022429418756029

[11] Elena Macrides and Charoula Angeli. 2018. Domain-Specific Aspects of Technological Pedagogical Content Knowledge: Music Education and the Importance of Affect. *TechTrends* (Jan 2018). https://doi.org/10.1007/s11528-017-0244-7

[12] Richard A. Magill. 2001. *Motor learning concepts and applications*. Boston, MA.

[13] Adnan Marquez-Borbon. 2016. Perceptual Learning and the Emergence of Performer-Instrument Interactions with Digital Music Systems. In *Proceedings of A Body of Knowledge - Embodied Cognition and the Arts conference (2016 Conference Proceedings)*. Paparazzi Press, Milan Italy, 422–431. https://doi.org/99.9999/woot07-S422

[14] Gongbing Shan Peter Visentin and Edwin B. Wasiak. 2008. Informing music teaching and learning using movement analysis technology. *International Journal Of Music Education* 26, 1 (2008), 73–87. https://doi.org/10.1177/0255761407085651

[15] B. Robson. 2004. Competition in sport, music and dance. *Medical Problems of Performing Artists* 19, 4 (2004), 160–165.

[16] Matthew D. Thibeault. 2018. Learning with Sound Recordings: A History of Suzuki's Mediated Pedagogy. *Journal of Research in Music Education* (2018), 0022429418756879. https://doi.org/10.1177/0022429418756879

[17] Merlin B. Thompson. 2018. What Does Music Learning Look like? In *Distributed Systems*. SpringerBriefs in Education, Springer, Cham, 19–34. https://doi.org/10.1007/978-3-319-65533-8_3

[18] Gualtiero Volpe, Paolo Alborno, Antonio Camurri, Paolo Coletta, Simone Ghisio, Maurizio Mancini, Radoslaw Niewiadomski, and Stefano Piana. 2016. Designing Multimodal Interactive Systems using EyesWeb XMI.. In *SERVE@ AVI*. 49–56.

[19] Gualtiero Volpe, Ksenia Kolykhalova, Erica Volta, Simone Ghisio, George Waddell, Paolo Alborno, Stefano Piana, Corrado Canepa, and Rafael Ramirez-Melendez. 2017. A multimodal corpus for technology-enhanced learning of violin playing. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter*. 25.

[20] Graham F. Welch. 1985. Variability of practice and knowledge of results as factors in learning to sing in tune. *Bulletin of the Council for Research in Music Education* (1985), 238–247.

[21] Clemens Wöllner, Frederik J.A. Deconinck, Jim Parkinson, and Michael J. Hove. 2012. The Perception of Prototypical Motion: Synchronization is Enhanced With Quantitatively Morphed Gestures of Musical Conductors. *American Psychological Association* 38, 6 (2012), 1390–1403.