BIVIB: A MULTIMODAL PIANO SAMPLE LIBRARY OF BINAURAL SOUNDS AND KEYBOARD VIBRATIONS

Stefano Papetti
Institute for Computer Music
and Sound Technology
Zürcher Hochschule der Künste
Zurich, Switzerland
`stefano.papetti@zhdk.ch`

Federico Avanzini
Dipartimento di Informatica
Università di Milano
Milan, Italy
`federico.avanzini@unimi.it`

Federico Fontana
Dipartimento di Scienze Matematiche,
Informatiche e Fisiche
Università di Udine
Udine, Italy
`federico.fontana@uniud.it`

ABSTRACT

An extensive piano sample library consisting of binaural sounds and keyboard vibration signals is made available through an open-access data repository. Samples were acquired with high-quality audio and vibration measurement equipment on two Yamaha Disklavier pianos (one grand and one upright model) by means of computer-controlled playback of each key at ten different MIDI velocity values. The nominal specifications of the equipment used in the acquisition chain are reported in a companion document, allowing researchers to calculate physical quantities (e.g., acoustic pressure, vibration acceleration) from the recordings. Also, project files are provided for straightforward playback in a free software sampler available for Windows and Mac OS systems. The library is especially suited for acoustic and vibration research on the piano, as well as for research on multimodal interaction with musical instruments.

1. INTRODUCTION

The multisensory aspects of musical performance have been studied since long, particularly focusing on sound and vibration [1, 2, 3, 4], and are recognized to have a major role in the complex perception-action mechanisms involved in musical instrument playing [5]. Indeed, during instrumental performance the musician is exposed to visual, haptic (i.e., tactile and kinesthetic), and of course auditory cues. Research in this direction has substantially gained momentum in recent years, as attested by the birth of new keywords such as “musical haptics” [6].

This increased interest is partly due to the availability of novel compact, accurate, and low-cost sensors and actuators, which enable the development of complex experimental settings for measuring and delivering multisensory information in real-time on a musical instrument during the performance [7, 8, 9, 10]. On the one hand these technologies offer the possibility to investigate the perceptual role of different sensory modalities in the interaction with traditional musical instruments, while on the other they enable the design of novel digital musical interfaces and instruments in which richer feedback modalities can increase the performer’s engagement, as well as the perceived quality and playability of the device [11, 12, 13, 14].

As a consequence, the availability of multimodal datasets combining and synchronizing different types of information (audio, video, MOCAP data of the instrument and the performer, physiological signals, etc.) is increasingly recognized as an essential asset for studying music performance and related aspects. Some recent examples include the “multimodal string quartet performance dataset” (QUARTET) [15], the “University of Rochester Multi-modal Music Performance dataset (URMP) [16], the “Database for Emotion Analysis using Physiological Signals” (DEAP) [17], as well as the RepoVizz initiative [18], which provides a system for storing, browsing, and visualizing synchronous multimodal data.

Within this general framework, the piano represents a relevant case study both for its prominence in the history of western musical tradition and for its potential in commercial applications (figures from the musical instrument industry[1] show a continuing growth of digital pianos and keyboard synthesizer sales).

When playing an acoustic piano, the performer is exposed to a variety of auditory, visual, somatosensory, and vibrotactile cues that combine and integrate to shape the pianist’s perception-action loop. The present authors are involved in a long-term research collaboration around this topic, with particular focus on the following two aspects. The first one is the tactile feedback produced by keyboard vibrations that reach the pianist’s fingers after keystrokes and holds until key release. The second one is the spatial auditory information contained in the sound field produced by the instrument at the performer’s head location. For both research fields, the existing literature is scarce and provides mixed if not contradictory results about the actual perceivability and possible relevance of this multisensory information [3]. We provide extensive discussion of these aspects in previously published studies, regarding both vibration perception [14] and sound localization [19] on the acoustic piano. Moreover, a digital piano prototype was recently developed that reproduces various types of vibrations [20] – including those recorded on acoustic pianos.

As part of this research, an extensive amount of experimental data has been produced during the past years. The purpose of this paper is to present an extensive multimodal piano sample library consisting of binaural sounds and keyboard vibration signals, some of which have been used in previous works for acoustic analysis and psychophysical testing, and has now been further expanded with upright piano data and organized into a single coherent open-access dataset. Section 2 presents the main features of the library, including a description of the hardware and software recording setups, and the organization of the samples for use in a free software sampler. Section 3 discusses some key aspects involved in the usage of the library, including sample analysis, multimodal playback, and several application scenarios.

---

[1] `https://www.namm.org/membership/global-report`

2. BUILDING OF THE BiVib SAMPLE LIBRARY

The BiVib (**Bi**naural and **Vib**ratory) sample library is a collection of high-resolution audio files (.wav format, 24-bit @ 96 kHz) representing binaural piano sounds and keyboard vibrations, accompanied by project files for a free software sampler, and documentation. The dataset, whose core structure is illustrated in Tab. 1, is made available through an open-access data repository[2] and released under a Creative Commons (CC BY-NC-SA 4.0) license.

2.1. Recording procedure

The samples were recorded on two Yamaha Disklavier pianos – a grand model DC3 M4 located in Padova, Italy, and an upright model DU1A with control unit DKC-850 located in Zurich, Switzerland. Disklaviers are MIDI-compliant acoustic pianos equipped with sensors for recording keystrokes and pedaling, and electromechanical motors for playback. The grand piano is located in a large laboratory space (approximately 6 × 4 m), while the upright piano is in an acoustically treated small room (approximately 4 × 2 m).

Recordings were acquired for 10 velocity values on each of the 88 keys by means of automated software-driven procedures sending MIDI messages, as described in detail further below.

2.1.1. Hardware setup

Binaural recordings made use of dummy heads with simulated ears and ear canals mounting binaural microphones, with slightly different setups for the grand and upright pianos: a system based on the KEMAR 45BM was used in Padova (PD), and a Neumann KU 100 in Zurich (ZH). The mannequins were placed in front of the pianos at the height and distance of an average pianist (see Fig. 1). The two binaural microphones were connected to the microphone inputs of two professional audio interfaces, respectively a RME Fireface 800 (PD, gain set to +40 dB) and a RME UCX (ZH, gain set to +20 dB). The condenser capsules of the microphones were respectively fed by 26CB preamplifiers powered by a 12AL power module (PD), and powered by 48 V phantom provided by the audio interface (ZH).

Three lid configurations were adopted for each piano. The grand piano (PD) was measured with the lid completely *closed*, completely *open*, and *removed* (i.e., physically detached from the main body of the piano). The upright piano was recorded with the lid *closed*, *semi-open* (see Fig. 1), and completely *open*. The purpose of using different configurations was to gain additional insight about the possible role of the lid in modulating the sound field reaching the performer's ears and related lateralization/localization cues [19]. As a result, three sets of binaural samples were recorded for each piano.

Vibration recordings were performed with a Wilcoxon Research 736 piezoelectric accelerometer connected to a Wilcoxon Research iT100M Intelligent Transmitter, whose AC-coupled output fed a line input of a RME Fireface 800 interface and was recorded as an audio signal. The accelerometer was manually attached with double-sided adhesive tape to each key in sequence, as depicted in Fig. 2.

Figure 1: *The binaural recording setup used in Zurich. The piano lid is in 'semi-open' position*

2.1.2. Software setup

Two different software setups were used respectively for sampling sound and vibration. The same MIDI velocity values were used in both cases: 10 values between 12 and 111, evenly spaced by 11-point intervals. This choice was based on a previous study by the present authors that determined a reliable range resulting in consistent acoustic intensity [14]: in fact, the electromechanical motors of computer-controlled pianos fall short – to different extent depending on the model – of providing a consistent dynamic response, especially for the lowest and highest velocity values [21].

Binaural samples were recorded via a fully automated procedure programmed in SuperCollider.[3] The recording sessions took place overnight, thus minimizing unwanted noise from personnel working in the building. On the grand piano, note durations were determined algorithmically, based upon their dynamics and pitch – ranging from 30 s used for A0 at velocity 111, to 10 s used for C8 at velocity 12 – so as to cover their full decay while minimizing the amount of recorded data and the length of recording session (still amounting to about 6 hours each). Indeed, notes of increasing pitch and/or decreasing dynamics have shorter decay times. Unfortunately, on the upright piano an undocumented protection mechanism prevents the electromechanical system from holding down the keys longer than about 17 s, thus not allowing to fully cover the notes' decay. Therefore, for the sake of simplicity all notes were recorded for just as long as possible.

Vibration samples were recorded through a slightly less so-

---

[2] `https://doi.org/10.5281/zenodo.1213210`

[3] A programming environment for sound processing and algorithmic composition: `http://supercollider.github.io/`.

Table 1: *Dataset core structure. Lid configurations used for binaural recordings are reported in square brackets*

	Disklavier DC3 M4 (grand, Padova)	**Disklavier DU1A with DKC-850 (upright, Zurich)**
Sample sets (.wav files)	Binaural [closed]	Binaural [closed]
	Binaural [open]	Binaural [semi-open]
	Binaural [removed]	Binaural [open]
	Keyboard vibration	Keyboard vibration
Sampler projects (Kontakt *multis*)	Binaural [closed] + vibration	Binaural [closed] + vibration
	Binaural [open] + vibration	Binaural [semi-open] + vibration
	Binaural [removed] + vibration	Binaural [open] + vibration

Figure 2: *The vibration recording setup: A Wilcoxon Research 736 accelerometer is attached with adhesive tape to a key that is being played remotely via MIDI control*

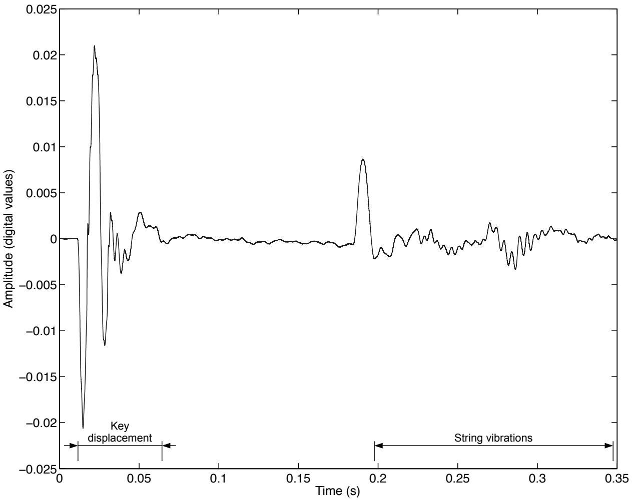

Figure 3: *Waveform of a vibration signal recorded on the grand Disklavier by playing the note A2 at MIDI velocity 12. Picture from [14]*

phisticated procedure. A DAW software was used to play back MIDI notes at the previously mentioned 10 velocity values while recording keyboard vibrations as audio signals. In this case, all notes had a fixed duration of 16 s that, considered the much weaker intensity of vibration signals as compared to sound, still allowed to describe the decay of vibration well beyond perceptual thresholds [14, 22].

2.2. Sample processing

Because of the intrinsic delay between sending MIDI messages from a computer and the mechanical actuation of the Disklavier pianos, the recorded samples started with a silent section, which we decided to remove especially in view of their use in a sampler (see 2.3). Given the large number of files (880 for each sample set), automated procedures were developed, tested and fine tuned, with the goal of removing the initial silence while leaving the rest unaffected.

Having been recorded through an accelerometer, vibration signals additionally had abrupt onsets in the attack, appearing in the first $200{-}250\,\mathrm{ms}$, and corresponding to the initial fly of the measured key followed by its impact with the piano keybed (see Fig. 3). As such, these onsets were not linked to sound-related vibratory cues at the keyboard, and therefore they had to be removed as well. Due the fact that onset profiles showed large variations, despite several tests made in MATLAB no reliable automated strategy could be found for editing the vibration samples. Therefore, a manual approach had to be employed instead: Files were imported in the Audacity sound editor, their waveform was zoomed in and auditioned, and the onset part was cut.

Sound recordings instead showed a more uniform shape, and an automated procedure programmed in SuperCollider was successfully used to cut the initial silence: For each sample, the program analyzes its amplitude envelope, detects the position of its largest peak, moves back by a few milliseconds, and finally applies a short fade-in.

2.3. Sampler projects and library organization

Project files are provided for use with the free ‘Player’ version of the software sampler Native Instruments Kontakt 5,[4] available for Windows and Mac OS systems. The full version of Kontakt 5 was instead used for developing the sampler projects. The library is organized into four folders named ‘Documentation’, ‘Instruments’,

---

[4] `https://www.native-instruments.com/en/products/komplete/samplers/kontakt-5-player/`

'Multis', and 'Samples'.

The 'Samples' folder – whose total size amounts to about 65 Gb – holds separate subfolders respectively for the binaural and vibration sample types, which in turn contain further subfolders for each sample set (see Table 1), for example 'grand-open' under the 'binaural' folder.

Independent of their type, sample files were named according to the following mask:

```
[note][octave #]_[lower MIDI velocity] ...
... _[upper MIDI velocity].wav
```

where `[note]` follows the English note-naming convention, `[octave #]` ranges from 0 to 8, `[lower MIDI velocity]` equals the MIDI velocity (range 12–111) used during recording and is the smaller velocity value mapped to that sample in Kontakt (see below), `[upper MIDI velocity]` is the greater velocity value mapped to that sample in Kontakt. For instance, a file `A4_100_110.wav` corresponds to the note A from the 4th octave (fundamental frequency 440 Hz) recorded at MIDI velocity 100, and mapped to the velocity range 100–110 in Kontakt. Since the lowest recorded velocity value was 12, no samples were mapped to the velocity range 1–11 in Kontakt.

Following Kontakt's terminology, each of the provided *instruments* reproduces a single sample set (e.g., binaural recording of the grand piano with lid open), while each *multi* combines two *instruments* respectively reproducing one binaural and one vibration sample set belonging to the same piano. The two *instruments* in each *multi* are configured so as to receive MIDI input data on channel 1, thus playing back at once, while their respective outputs are routed to different virtual channels in Kontakt: binaural samples are routed to a pair of stereo channels (numbered 1-2), while vibration samples are played through a mono channel (numbered 3). In this way, when using audio interfaces offering more than two physical outputs, it is possible to render both binaural and vibrotactile cues at the same time by routing the audio signal respectively to headphones and vibration actuators.

In each *instrument*, sample mapping was implemented relying on the 'auto-map' feature found in the full version of Kontakt: this parses file names and uses the recognized tokens for assigning samples to e.g. a pitch and velocity range. The chosen file naming template made it straightforward to batch-import the samples.

The amplitude of the recorded signals was not altered, that is no dynamic processing or amplitude normalization was applied, and the volume of all Kontakt *instruments* was set to 0 dB. Because of this and the adopted velocity mapping strategy, sample playback is made transparent for acoustic and vibratory analysis and experiments (see 3.1 and 3.2).

3. USING THE BiVib SAMPLE LIBRARY

The BiVib library is suited for both acoustic/vibratory analysis and interactive applications, for instance in experiments on musical performance and multisensory perception.

To our knowledge, no other existing piano datasets are fully comparable with what included with the BiVib library. Indeed, binaural piano sounds are offered by a few audio plugin developers (e.g., Modartt Pianoteq[5]) and digital piano manufacturers (e.g., Yamaha Clavinova[6]). Also, free binaural piano samples can be found, such as the "binaural upright piano" library,[7] which however offers only 3 dynamic layers as opposed to the 10 velocity levels provided by BiVib. Overall, such binaural sounds are conceived for use with virtual instruments, while they are not directly suitable for research purposes, due to non-reproducible and undocumented acquisition procedures and sample post-processing. Collections of haptic / vibrotactile data of musical instruments are even scarcer. To our knowledge, no other public dataset of piano keyboard vibrations is available.

3.1. Sample analysis

For many experimental purposes and applications it is essential to be able to reconstruct the physical values of the measured signals, that is acceleration in $\mathrm{m/s^2}$ for keyboard vibrations, and acoustic pressure in Pa for the binaural signals. Given the quality of the equipment used in the various stages of the acquisition chain, such reconstruction can be achieved with good accuracy by relying on the equipment's nominal specifications. These are summarized in a companion document included in the 'Documentation' folder.

For instance, accelerations in $\mathrm{m/s^2}$ can be computed from the acquired signals by making use of the nominal sensitivity parameters of the audio interface and the accelerometer: the digital signals, whose normalized values range between -1 and 1, are first converted to voltage values through the full scale reference of the RME Fireface 800 audio interface (for line inputs at the chosen sensitivity level, $0\,\mathrm{dBFS}$ @ $+19\,\mathrm{dBu}$, reference $0.775\,\mathrm{V}$), and then transformed into proportional acceleration values through the sensitivity constant of the Wilcoxon Research 736 accelerometer ($10.2\,\mathrm{mV/m/s^2}$). In a similar way, acoustic pressure values in Pa can be obtained from the binaural recordings, by making use of the nominal sensitivity levels of the audio interfaces' microphone inputs and of the binaural microphones.

Generally speaking, objective data computed from the library may help support results from psychophysical and quality evaluation studies focusing on the piano, as recently done by the authors in [14].

A more ambitious task could be that of extracting piano sounds free of the room response that affect the BiVib library. Methods exist to deconvolve common acoustic poles and zeros from samples that have been captured under invariant conditions [23], as it is in our case. However, in the case of BiVib care should be taken for preventing these methods from cancelling poles and zeros that are introduced by the mannequin, responsible of the binaural cues: Most such poles and zeros have frequencies higher than those associated to the dominant poles and zeros characterizing the recording rooms, in ways that at least the lower common modal resonances may be deconvolved safely from the samples. On the other hand, anechoic binaural sounds may not be suitable for the purpose of listening experiments in ecological settings.

3.2. Experiments and applications

We anticipate that this library will be useful for data analysis and experiments in music performance studies.

Acceleration values in $\mathrm{m/s^2}$ obtained from the vibration recordings as explained above can be used e.g. for comparison with the literature of touch psychophysics [22, 24], as shown in Fig. 4. In a recent article by the present authors, this allowed to

---

[5] `https://www.pianoteq.com/`

[6] `https://europe.yamaha.com/en/products/musical_instruments/pianos/clavinova/`

[7] `https://www.michaelpichermusic.com/binaural-upright-piano`

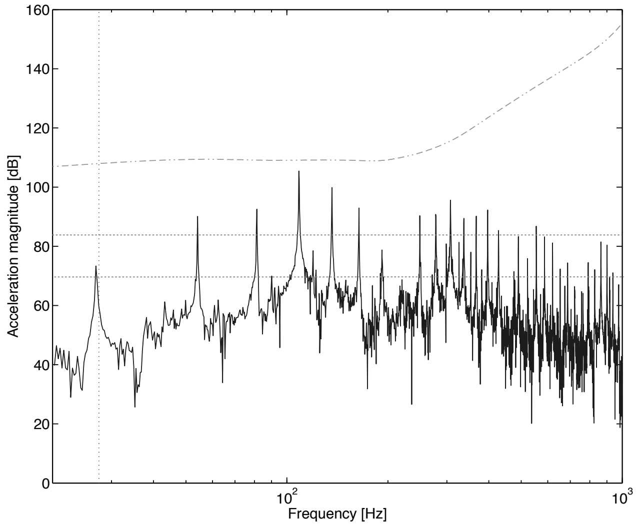

Figure 4: *Magnitude spectrum of the vibration signal at the A0 key, recorded with MIDI velocity 111 on the upright Disklavier. The dash-dotted curve depicts the reference vibrotactile threshold for passive touch [24], while the two horizontal dashed lines represent the minimum and maximum thresholds recently measured by one of the authors for active touch [22]. Picture adapted from [14]*

support the subjective results of a psychophysical experiment on the detection of vibration at the piano keyboard [14].

On a genuinely multisensory level, the relations in intensity existing between sound and vibration signals, recorded on the same instruments and provided by the database, may be used to investigate the presence of cross-modal effects occurring during piano playing. Such effects have been highlighted as part of a more general multisensory integration mechanism [25] that under certain conditions may increase the perceived intensity of auditory signals [26], or vice-versa can enhance touch perception [27]. The possibility to individually manipulate the magnitude of piano sounds and vibrations in experimental settings (e.g., using a digital keyboard that yields multimodal feedback) may lead to interesting observations on the perceptual consequence of this manipulation specifically for the pianist. In this regard, cross-modal effects resulting from varying the tactile feedback of the keyboard have been recently observed by the authors, however far from giving a systematic view about the impact of the different sensory channels to the pianist's playing experience [20].

The BiVib library has been previously used to investigate the presence of auditory lateralization cues for the acoustic piano, limited to sound samples. Although the recordings are not anechoic, their reproduction through headphones has unveiled the ability of pianists to localize tones in good accordance with the interaural level differences existing in the binaural material [28]. This ability was further supported by visual cues of self-moving keys producing the corresponding tones, as well as by somatosensory cues occurring during active piano playing of the same tones [19]. Interestingly, the supportive role of the visual and somatosensory channel ceased when the auditory feedback was subverted by swapping the left-right signals feeding the headphones. This evidence speaks in favor of the existence of a ventriloquist effect that affects piano listening and playing, which may be enabled only by a coherent multisensory experience as provided by an actuated piano [28].

One promising research direction that may also gain from using the BiVib library is represented by the use of methods from cognitive neuroscience (e.g., EEG and event-related potentials, brain imaging) to further investigate the role of multimodal audio-visuo-tactile processing in supporting musical abilities and triggering the activation of motor information in the brain of pianists.

Ultimately, all these studies can contribute to the perceptually and cognitively informed design of novel digital pianos, and to the understanding of perceived instrumental quality and playability. We provided initial results in an earlier study where we developed and tested a haptic digital piano prototype: various vibration signals, including grand piano vibrations from BiVib, were reproduced at the keyboard and compared to a non-vibrating condition [20]. Overall, vibrating condition was preferred over the standard non-vibrating setup in terms of perceived quality. However, when considering performance-related features such as timing and dynamics accuracy of performers, this initial study could not highlight significant differences between conditions.

Finally, the binaural recordings may be especially useful also for different research directions. One example in the field of music information retrieval is that of multipitch estimation and automatic transcription algorithms that exploit binaural information, whereas the datasets most commonly employed for these tasks are not binaural, such as the "MIDI Aligned Piano Sounds" (MAPS) database [29]. One further example, in the field of digital audio effects, is that of spatial enhancement effects (e.g., stereo enhancement): Piano sounds are typical examples of acoustic signals that are difficult to spatialize properly [30], and the BiVib samples may serve as a reference for the development/validation of novel effects.

4. CONCLUSIONS AND PERSPECTIVES

The BiVib sample library provides a unique set of multimodal piano data, acquired with high-quality equipment in controlled conditions through reproducible computer-controlled procedures.

Since the binaural samples in the library were meant for use in perceptual tests under ecological listening conditions, they currently include responses of the rooms where they were recorded. However we recognize that for acoustic research purposes this may be a relevant limitation, and therefore we have planned to add the respective (binaural) room impulse responses in a future version of the library, and possibly a complete new set of recordings in anechoic conditions.

We hope that the public availability of the library, in conjunction with this documentation and with the accompanying Kontakt sampler projects, will facilitate further research in the understanding and modeling of piano acoustics, performance, and related fields.

5. ACKNOWLEDGMENTS

This research was partially supported by project AHMI (Audio-haptic modalities in musical interfaces, 2014–2016), and HAPTEEV (Haptic technology and evaluation for digital musical interfaces 2018–2022), both funded by the Swiss National Science Foundation.

The Disklavier grand model DC3 M4 located in Padova was made available by virtue of the Sound and Music Processing Lab

(SaMPL), a project of the Conservatory of Padova funded by Cariparo foundation (thanks in particular to Nicola Bernardini and Giorgio Klauer).

The authors would like to thank several students and collaborators who contributed to the development of this work along the years, in chronological order: Francesco Zanini, Valerio Zanini, Andrea Ghirotto, Devid Bianco, Lorenzo Malavolta, Debora Scappin, Mattia Bernardi, Francesca Minchio, Martin Fröhlich.

6. REFERENCES

[1] K. Marshall and B. Genter, "The musician and the vibrational behavior of a violin," *J. of the Catgut Acoustical Society*, vol. 45, pp. 28–33, 1986.

[2] H. Suzuki, "Vibration and sound radiation of a piano soundboard," *J. Acoust. Soc. Am.*, vol. 80, no. 6, pp. 1573–1582, 1986.

[3] A. Askenfelt and E. V. Jansson, "On vibration sensation and finger touch in stringed instrument playing," *Music Percept.*, vol. 9, no. 3, pp. pp. 311–349, 1992.

[4] C. Saitis, "Evaluating violin quality: Player reliability and verbalization," Ph.D. dissertation, Dept. of Music Research, McGill University, Montreal, Canada, 2013.

[5] S. O'Modhrain and B. R. Gillespie, "Once more, with feeling: The dynamics of performer-instrument interaction," in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Springer-Verlag, 2018, in press.

[6] S. Papetti and C. Saitis, Eds., *Musical Haptics*. Springer-Verlag, 2018, in press.

[7] S. O'Modhrain and C. Chafe, "Incorporating Haptic Feedback into Interfaces for Music Applications," in *Proc. of ISORA, World Automation Conf.*, 2000.

[8] M. T. Marshall and M. M. Wanderley, "Vibrotactile feedback in digital musical instruments," in *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, 2006, pp. 226–229.

[9] D. M. Birnbaum and M. M. Wanderley, "A systematic approach to musical vibrotactile feedback," in *Proc. Int. Computer Music Conf. (ICMC)*, 2007.

[10] D. Overholt, E. Berdahl, and R. Hamilton, "Advancements in actuated musical instruments," *Organised Sound*, vol. 16, no. 02, pp. 154–165, 2011.

[11] M. Keane and G. Dodd, "Subjective Assessment of Upright Piano Key Vibrations," *Acta Acust. united with Acust.*, vol. 97, no. 4, pp. 708–713, 2011.

[12] A. Galembo and A. Askenfelt, "Quality assessment of musical instruments - Effects of multimodality," in *Proc. Conf. of the European Society for the Cognitive Sciences of Music (ESCOM)*, Hannover, Germany, Sep 2003.

[13] I. Wollman, C. Fritz, and J. Poitevineau, "Influence of vibrotactile feedback on some perceptual features of violins," *J. Acoust. Soc. Am.*, vol. 136, no. 2, pp. 910–921, 2014.

[14] F. Fontana, S. Papetti, H. Järveläinen, and F. Avanzini, "Detection of keyboard vibrations and effects on perceived piano quality," *J. Acoust. Soc. Am.*, vol. 142, no. 5, pp. 2953–67, 2017.

[15] E. Maestre, P. Papiotis, M. Marchini, Q. Llimona, O. Mayor, A. Pérez, and M. M. Wanderley, "Enriched multimodal representations of music performances: Online access and visualization," *IEEE MultiMedia*, vol. 24, no. 1, pp. 24–34, 2017.

[16] B. Li, X. Liu, K. Dinesh, Z. Duan, and G. Sharma, "Creating a multi-track classical music performance dataset for multi-modal music analysis: Challenges, insights, and applications," *IEEE Trans. Multimedia*, 2018, submitted for publication.

[17] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis; using physiological signals," *IEEE Trans. Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.

[18] O. Mayor, J. Llop, and E. Maestre Gómez, "Repovizz: A multi-modal on-line database and browsing tool for music performance research," in *Proc Int. Soc. for Music Information Retrieval Conf. (ISMIR 2011)*, Oct. 2011.

[19] F. Fontana, D. Scappin, F. Avanzini, M. Bernardi, D. Bianco, and G. Klauer, "Auditory, visual and somatosensory localization of piano tones: A preliminary study," in *Proc. Int. Conf. Sound and Music Computing (SMC)*, Espoo, Jul. 2017, pp. 254–260.

[20] F. Fontana, F. Avanzini, H. Järveläinen, S. Papetti, G. Klauer, and L. Malavolta, "Rendering and subjective evaluation of real vs. synthetic vibrotactile cues on a digital piano keyboard," in *Proc. Int. Conf. Sound and Music Computing (SMC)*, Maynooth, Ireland, Jul. 2015, pp. 161–167.

[21] W. Goebl and R. Bresin, "Measurement and reproduction accuracy of computer-controlled grand pianos," *J. Acoust. Soc. Am.*, vol. 114, no. 4, pp. 2273–83, 2003.

[22] S. Papetti, H. Järveläinen, B. L. Giordano, S. Schiesser, and M. Fröhlich, "Vibrotactile sensitivity in active touch: effect of pressing force," *IEEE Trans. on Haptics*, vol. 10, no. 1, pp. 113–122, Jan 2017.

[23] Y. Haneda, S. Makino, and Y. Kaneda, "Common Acoustical Pole and Zero Modeling of Room Transfer Functions," *IEEE Trans. on Speech and Audio Processing*, vol. 2, no. 2, pp. 320–328, Apr 1994.

[24] R. T. Verrillo, "Vibration sensation in humans," *Music Percept.*, vol. 9, no. 3, pp. 281–302, 1992.

[25] C. Kayser, C. I. Petkov, M. Augath, and N. K. Logothetis, "Integration of Touch and Sound in Auditory Cortex," *Neuron*, vol. 48, no. 2, pp. 373–84, oct 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16242415

[26] H. Gillmeister and M. Eimer, "Tactile enhancement of auditory detection and perceived loudness," *Brain Research*, vol. 1160, pp. 58 – 68, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0006899307006671

[27] T. Ro, J. Hsu, N. E. Yasar, L. C. Elmore, and M. S. Beauchamp, "Sound enhances touch perception," *Experimental Brain Research*, vol. 195, pp. 135–143, 2009.

[28] F. Fontana, F. Avanzini, and S. Papetti, "Evidence of lateralization cues in grand and upright piano sounds," in *Proc. Int. Conf. Sound and Music Computing (SMC)*, Cyprus, 2018, submitted.

[29] V. Emiya, R. Badeau, and B. David, “Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle,” *IEEE Trans. Audio, Speech, and Language Processs.*, vol. 18, no. 6, pp. 1643–1654, 2010.

[30] D. Rocchesso, “Spatial effects,” in *Digital Audio Effects*, U. Zölzer, Ed. Chirchester Sussex, UK: John Wiley & Sons, 2002, pp. 137–200.

[31] S. Soto-Faraco and G. Deco, “Multisensory contributions to the perception of vibrotactile events,” *Behavioural Brain Research*, vol. 196, no. 2, pp. 145–154, 2009.