Hand Gestures in Music Production

Axel Berndt, Simon Waloschek, Aristotelis Hadjakos
Center of Music and Film Informatics
University of Music Detmold, Detmold, Germany
{berndt|waloschek|hadjakos}@hfm-detmold.de

ABSTRACT

Today's music production interfaces are still dominated by traditional concepts. Music studios are a prime example of a multi-device environment with some of the most complex user interfaces. In this paper, we aim at introducing a more embodied interaction modality, freehand interaction, to this scenario. Not as replacement! We analyze typical music production scenarios and derive where freehand input yields added value. We give an overview of related research and discuss aspects of the technical integration into music studio setups. This is complemented by a prototype implementation and a survey that provides clues for the development of an intuitive gesture set for expressive music control.

1. MOTIVATION

Music studios were always multi-device environments with specialized equipment for literally every task, such as sound synthesis, audio effects, recording, and mixing. At the center of this network lies the *Digital Audio Workstation* (DAW), a computer that integrates all hardware components and software supplements. Despite this high degree of specialization and distribution of functionality over a multitude of networked components we still observe a predominant uniformity of user interface concepts which are rooted in classic hardware devices. While new interface technologies are vitally incorporated in the context of digital musical instruments, we cannot register a similar fertilization in the field of music production. One main reason for this is a considerable interference with established optimized workflows. While each developer in this field is confronted with this problem, their thoughts and solutions were neither sufficiently documented nor discussed so far.

This paper addresses freehand gesture input. Our aim is not an entire replacement of established input modalities. Knobs, sliders, keyboard and mouse proved their worth and can be regarded as optimal solutions for many tasks. We want to keep established workflows intact. The greatest gain of freehand input lies in the continuous control of big ranges of parameters that develop over time, live with the musical playback. The control data that derives from the hand gestures can be converted to MIDI data. If music is produced solely electronically, without recorded human musicians, this may be used to steer expressive parameters such as tempo, dynamics, articulation, and timbre. This is our target scenario in this text. First, we give an overview of exemplary previous and related work in this field. The technical integration into the typical DAW setup is discussed in section 3. In section 4 we report of a survey to gain first cues for developing an intuitive set of gestures.

2. HANDS ON MUSIC: RELATED WORK

Many approaches to freehand gesture-controlled expression are based on the tracking of hand-held devices such as batons [1], drumsticks [2], and balls [3]. These may not only be tracked externally but may be equipped with sensors themselves, such as the Wii Remote, smartphones [4] and others [5, 6]. To avoid the necessity of holding a device, sensors can be fixed at the hand [7] or data gloves can be used [8, 9]. With optical tracking systems such as the HandSonor system [10], Microsoft's Kinect, and the Leap Motion no hand-held devices are necessary at all.

Typical software frameworks for musical freehand interaction are the commercial GECO system [11] and the Harmonic Motion toolkit [12]. The authors of the latter performed a preparatory survey of potential users. One outcome was the fact that 57% of their participants saw the most problematic issue of gestural music interaction in the mapping of gesture data to musical data. Speed, latency, hardware-specific problems, and stability came far after.

Wanderley [13] lists five different musical contexts of gestural control in music of which the following three are directly related to the process of music making and production and will be in the focus of the subsequent overview of related work in the field. *Instrument manipulation* takes place at the level of realtime sound synthesis control of digital musical instruments. *Score-level control*, such as conducting, manipulates the expression of a performance, not the material performed. *Post-production activities* address the control of digital audio effects, sound spatialization, and mixing. The large corpus of works in this field cannot be recapitulated completely here. We will pinpoint representative and illustrative works. A more comprehensive treatment of the subject can be found in [14, 15].

Sridhar [10] distinguishes continuous and discrete gesture data which can be mapped to likewise continuous and discrete instrument parameters. Dahl [16] focusses on the latter, when he studies air drumming gestures to identify motion features that provide the best timing information for discrete musical triggering. Françoise et al. [17] bypass the necessity of an explicit definition of the gesture-to-sound mapping by a machine learning method.

Gossmann & Neupert [18] introduce the *remix instru-*

ment. Recorded audio material of an instrument is analyzed and its atomic sounds are arranged in a 3d scatter plot. The user's hand moves the playback cursor of a concatenative synthesis through this 3d space. The artistic installation, "Non Human Device #002" [19], allows for the control of sound parameters through freehand interaction with a jellyfish-like virtual creature. Two further instruments, the Air-Keys and the Air-Pads [20], are critically discussed in a lessons-learned report giving practical hints on playability and tracking performance for such projects. Further mappings for effects and synthesis modulation are described by Hantrakul & Kaczmarek [21].

The VUZIK/ChoirMob interface [22] performs a predefined polyphonic score. The performers manipulate the expression of the synthesized vocal sounds via touch gestures on smartphones. This work demonstrates the coupling of predefined musical "raw material" and its expressive realtime manipulation through gestures. Such predefined musical material does not necessarily have to be static but can also be generated in realtime. Tormoen et al. [23] use the Leap Motion controller in combination with the Rubato Composer for gesture-based music composition and improvisation, with particular regard to the transformation of musical material. Such interactive, semi-automatic composing and performance systems constitute a seamless fade between instrument- and score-level interaction.

Hand tracking-based music conducting systems represent a popular type of score-level applications. Lim & Yeo [4] track conducting gestures via the gyroscope sensor of a smartphone. The MICON system [24] is an interactive exhibit that optically tracks a baton with an infrared sensor.

Gesture-controlled audio mixing is another recurrent subject [7, 25, 26]. Balin & Loviscach [27] see the chance of reducing the complexity of traditional DAW's GUIs via gestural control elements. They developed and evaluated a mouse and touch-operated gesture set. Ratcliffe [28] visualizes audio tracks as spheres in a 3d space that can be grasped for stereo positioning and depth mixing. A similar so-called *stage metaphor* has been adopted by Lech & Kostek [29] who further propose a comprehensive set of symbolic hand gestures. This may further help to alleviate the attention dragging, adulterating visuality of traditional DAW GUIs [29, 30].

All these works show freehand gesture control being an interaction modality that holds several promising perspectives for music production beyond the pure instrument playing. This requires both, its introduction to established workflows and the development of appropriate gesture sets, including corresponding mappings. Not all functionality in music production benefits from this input type. For some functions faders, knobs, mouse, and keyboard are an optimal choice. Freehand gestures should rather be seen as a complement than a replacement. The following section puts the focus on the technical side and discusses the integration of freehand gesture control in typical DAW setups.

3. HANDS ON THE DAW

This section discusses several noteworthy aspects of the technical integration of freehand interaction into a typical DAW environment. Figure 1 gives an overview of the resulting architecture.

Several hand tracking solutions are available, today. Very popular is Microsoft's Kinect 2. Its comparatively low resolution and frame rate of about 30fps, however, make it more suitable for expansive whole-body gestures and a rather rough control of musical parameters, e.g. slow and steady dynamics and tempo changes. Fast fine-grained control, such as note-wise articulations, are impractical. The Leap Motion controller, on the other hand, is specialized for hand tracking. Its tracking range is distinctly smaller compared to the Kinect but offers a superior resolution that allows for very fine-grained hand poses and gestures. Even very fast gestures can be detected reliably thanks to its sampling rate of up to 300fps. Since it was designed for the use on tables, it seems the device of choice for professional audio workstations and is small enough to easily find a place in-between the other devices which is a key advantage over many other tracking systems that involve several cameras distributed around the tracking area. The user can quickly switch between gestural input and any other device. Especially when using the non-dominant hand for gesture input while keeping the dominant hand on a primary device seems advantageous in this scenario.

Next, a gesture recognizer identifies meaningful gestures. A mapping process converts these into a format that modern DAWs can further process and record, e.g. MIDI. Inside the DAW, the MIDI data can be used to control various parameters of sound synthesis and effect plugins as well as the overall mix. Frameworks such as the commercial GECO system [11] allow users to define their own mappings. For controlling solely sound-related parameters of a given musical material, the software chain ends here. However, this does not allow for more complex control of tempo, micro timing or alteration of the note material. More sophisticated tasks, e.g. gesture-driven realtime algorithmic music generation, require an additional, more flexible MIDI sequencer. In such a case, the DAW is only used as a sound and effects generator and recording device. The standalone sequencer takes over the responsibility to read, modify and even generate MIDI data.

Direct feedback during input generally eases the process of learning new input modalities and reduces users' mental load. The user should get notified about not or wrongly detected gestures instead of being frustrated by opaque decisions of the gesture recognizing system. Therefore, the visualization of tracking data (body, hands, depth sensor, gestures) as well as the audio output and additional auditory cues, presented in realtime, are advisable and allow for quick reactions and adaptations by the user. Such feedback requires a low latency to the gestural input. This requirement may be relaxed in non-live situations where no discrete and precisely timed sound events have to be entered.

4. SURVEY ON GESTURES

As we have pointed out previously freehand interaction in music production scenarios has, in our opinion, its greatest potential in the control of musical parameters that are otherwise laborious to handle, in particular multiple continuous parameters at once and live with the musical playback. Typical candidates for this are sound synthesis and audio effects parameters as well as expressive performance

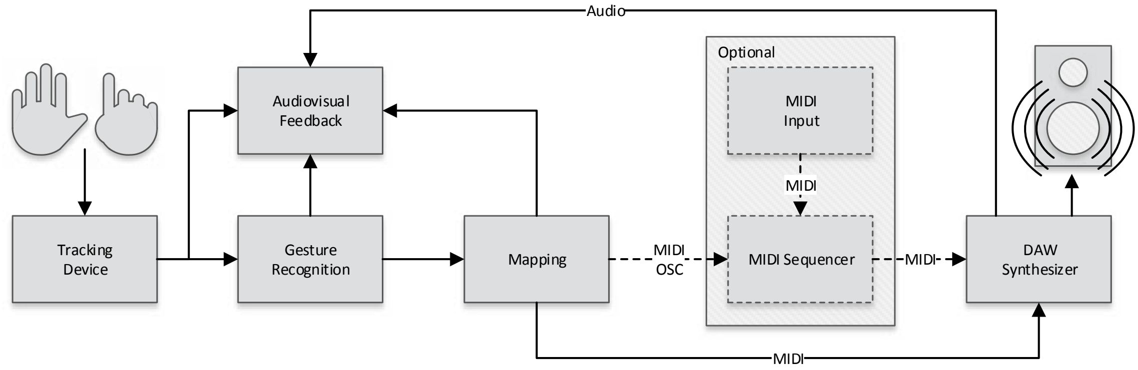

Figure 1. Integration of freehand interaction into a DAW.

parameters (assuming that the music is partly or solely produced on the computer). The range of such, possibly interesting parameters is wide. The concept of digital musical instruments is related to the application scenarios addressed but additionally involve the triggering of musical events. In contrast to this, we regard the musical raw material, i.e. the notes, as fixed, but not the way they are performed. Multi-channel mixing and conducting have already been addressed by many others. Hence, we decided to focus on the expressive performance, i.e. timing, dynamics, articulation, and timbre, which were considered only rarely in previous work so far.

In search of an intuitive gesture set we conducted a survey. The participants were asked to suggest gestures they find intuitive to manipulate tempo (faster/slower), dynamics (louder/softer), articulation (legato/staccato), and timbre (brighter/darker or shriller/softer). This covers only a subset of the manifold possibilities to manipulate expressive music performances which cannot all be included in only one single survey. Hence, we decided to focus on the most prominent features first and see if the supposed gestures may already be applicable also to more fine-grained features such as metrical accentuation and rubato. A follow-up survey can then focus on these and invite participants with a respectively more professional background.

4.1 Setup, Participants, Conduct

The survey took place during an open house event at Ostwestfalen-Lippe University of Applied Sciences at May 9th 2015 in Lemgo, Germany from 10am to 4pm. The location was a seminar room (doors kept open) equipped with speakers and projector. These were used to operate a prototype implementation. Here the visitors got used to the Leap Motion controller and could produce sound output with hand gestures. Besides the more playful wind and laser sword-like sounds the users could manipulate live-generated music (homophonic chorale) according to the four parameters of the survey. The predefined one-hand gestures—actually poses—were as follows:

Tempo was controlled with the depth position of the hand. Moving the hand forward increases the musical tempo, putting the hand back slows the tempo down.

Dynamics derived from the vertical hand position. Moving the hand up and down causes the volume level to increase and decrease.

Articulation was controlled by the grab strength. The flat hand produced a legato and the fist a staccato articulation. For poses in-between the note lengths were interpolated.

Timbre manipulation was done by hand tilting between the horizontal (soft sound) and vertical (shrill sound) pose. The timbres were achieved by oscillator hard syncing.

All parameters are controlled at once. The horizontal axis was left unused to be able to take the hand out of the tracking frustum and keep a certain parameter setting.

Among the numerous visitors, 44 took part in the interview (23 male, 21 female) aged from preschool to retirement age. None of them had a professional music background, hence there was few bias towards classic conducting gestures. Some participants got to know the prototype demo already before the interview. In these cases their answers could have been biased by the predefined gestures. In case that they answered these same gestures we insisted in further suggestions and counted the demo gesture suggestion in the evaluation only if the participants still explicitly preferred it over the own suggestions. The interviews were video recorded from the front. In the evaluation we collected all gestures and counted the suggestions.

4.2 Results & Discussion

We collected 281 suggestions of 115 gestures including those that are repeatedly suggested for different tasks. In some cases, the participants suggested gestures that indicate a certain parameter value, e.g. tempo specification by beat knocking. As this implicitly includes increase and decrease, e.g. knocking faster or slower, we counted these suggestions twice, i.e. once for increase and once for decrease. We identified 47 gesture pairs, i.e., two identical but inversely performed gestures.

Tempo Gestures: For tempo control we had 71 suggestions of 31 gestures. Only 14 suggestions (19.7%) of 10 gestures involved both hands. 26 suggestions (36.6%) of 9 gestures were actually poses that specify the tempo, e.g., through grab strength or the vertical position of the hand. All others indicate the tempo through their motion characteristics. The demo gesture (quasi a forward/ backward lever) was never suggested. The top-rated gestures are shown in table 1.

Further variants of the "calm-down gesture" $t1$ differed by using one or two hands, orienting the palm generally

Accelerando (increase tempo, 100% = 29 suggestions)

$T1$	fast fanning away with the back of one (open) hand, also described as wiping away (9 suggestions, 31%)
$T2$	fast circular movement of one (open) hand, away from the body, back of the hand heading forward (5 suggestions, 17.2%)
$T3$	one hand moves to the right, also described as fast-forwarding on a video recorder (4 suggestions, 13.8%)

Ritardando (decrease tempo, 100% = 34 suggestions)

$t1$	one (open) hand, palm downward, moves downward, also described as "calm-down gesture" (11 suggestions, 32.4%)
$t2$	opposite of gesture $T3$, one hand moves to the left, also described as rewinding on a video recorder (4 suggestions, 11.8%)

Table 1. Suggested gestures for musical tempo control.

Crescendo (increase volume level, 100% = 45 suggestions)

$D1$	one (open) hand, palm downward, moves upward (17 suggestions, 37.8%)
$D2$	both hands open, palms facing each other, spreading arms. In some cases the movement triggers the crescendo irrespective of the distance of both hands. Other participants expressed a specific loudness value through the distance between both hands. (5 suggestions, 11.1%)
$D3$	similar to $D1$ but with palm heading upward (4 suggestions, 8.9%)
$D4$	similar to $D1$ but with two hands (4 suggestions, 8.9%)
$D5$	similar to $D1$ but with two hands and palm heading upward (4 suggestions, 8.9%)
$D6$	one (open) hand, held vertical, moves to the right (3 suggestions, 6.7%)

Decrescendo (decrease volume level, 100% = 48 suggestions)

$d1$	opposite of gesture $D1$, one (open) hand, palm downward, moves downward (18 suggestions, 37.5%)
$d2$	opposite of gesture $D4$, similar to $d1$ but with both hands (16 suggestions, 33.3%)
$d3$	opposite of gesture $D2$, both hands open, palms facing each other, bringing arms together (3 suggestions, 6.3%)
$d4$	opposite of gesture $D6$, one (open) hand, held vertical, moves to the left (3 suggestions, 6.3%)

Table 2. Suggested gestures for musical dynamics control.

downward or into the direction of movement and use of a decent shaking to activate or intensify the input. In all variants the downward movement of the hands was used for slowdown and upward movement for acceleration. Even though each variant was suggested only once or twice, altogether (including the above "calm-down gesture") we had 18 suggestions (25.4% of all tempo suggestions).

We had also 18 suggestions (25.4% of all tempo suggestions) of "beat knocking" gestures (up and down movement) in different variants, including one or two symmetrical moving hands, open hand or stretched index finger, hand clapping, and other rhythmic hand motion. Here, the tempo derives from the pace of the motion pattern.

Dynamics Gestures: For dynamics control the participants had 93 suggestions of 25 gestures, including 81 suggestions (87.1%) of 19 poses and 36 suggestions (38.7%) of 9 bimanual gestures. This reflects a preference of one-handed poses for dynamics control. Table 2 shows the top-rated dynamics gestures.

Most of these gestures are variants of gesture $D1$ and $d1$, respectively. In sum, we got 65 suggestions (70% of all dynamics suggestions) for vertical hand movement, upward to increase and downward to decrease the volume level. These were already implemented in the demo and, as far as the participants knew them, widely confirmed.

Articulation Gestures: The participants gave 66 suggestions of altogether 21 gestures. This includes 23 suggestions (34.8%) of 10 poses and 25 suggestions (37.9%) of 7 bimanual gestures (top-rated gestures in table 3).

Gestures $a1$ and $a2$ were always jointly suggested. Their only difference is the use of one or both hands. Thus, we see them as equivalent. We further observed that the participants preferred gestures that involve grabbing/finger spreading and smooth/choppy movements to indicate articulation (53 of 66 suggestions, 80.3%).

Timbre Gestures: This musical feature was perceived as the most difficult to express through gestures. This mirrors not only in many oral comments but also in a greater diversity of suggested gestures (all together 51 suggestions of 38 gestures) and a low maximum score of 4. We collected 36 suggestions (70.6%) of 27 poses and 15 suggestions (29.4%) of 12 bimanual gestures. This indicates one-handed poses to be the preferred gesture types to control timbre. The top-rated gestures are listed in table 4.

22 suggestions (34.1%) specifically involved the fingers in some way, be it in the form of fast, chaotic or slow, wavy finger movement or the fingers' position (claw-like, flat, spread, or in right angle with the palm). Such a variety was not observed for the other musical parameters.

Discussion: Although we asked those participants who suggested gestures from the demo to make further suggestions and think about what they find intuitive, bias cannot entirely be excluded. On the other side, some of the demo gestures were not even mentioned or only once (specifically the gestures for tempo and timbre control). This fact suggests that the bias was not strong and/or the corresponding demo gestures not very successful.

We generally observed a preference of one-handed gestures. Only 90 suggestions (13.5%) out of 281 involved both hands. Regarding the typical music production workstation, where the user sits at a computer or mixer, the one-handed input is advantageous. Here, the user can keep one hand at the primary device and make freehand input with the secondary hand "by the way". This is also a good starting point for introducing multimodal interaction concepts.

Our results also include concurrent gestures, i.e. similar gestures for different tasks (e.g., $t1 = d1 = s3$, $T3 = D6$, $t2 = d4$, and $S3 = D1$). Hence, gesture combinations for parallel control of all four parameters are not possible with only the top-rated gestures. Instead, we will have to find a

Legato (broad articulation, 100% = 24 suggestions)

$A1$	smooth horizontal wave movement of one open hand, palm heading downward (6 suggestions, 25%)
$A2$	both hands open and arms wide open, also described as indicating a long tone length (5 suggestions, 20.8%)
$A3$	one hand, open/flat (4 suggestions, 16.7%)

Staccato (short articulation, 100% = 42 suggestions)

$a1$	rhythmic dabbing movement with fist, beak pose or thumb and index finger with one hand (11 suggestions, 26.2%)
$a2$	similar to $a1$ but with both hands moving symmetrical (11 suggestions, 26.2%)
$a3$	opposite of $A2$, both hands open, held close to each other, also described as indicating a short tone (4 suggestions, 9.5%)
$a4$	opposite of $A3$, fist (4 suggestions, 9.5%)
$a5$	opposite of $A1$, one (open) hand, vertically held, makes choppy up and down movements, also described as hacking with the side of the hand (4 suggestions, 9.5%)

Table 3. Suggested gestures for musical articulation control.

good trade-off in our further steps. An approach might be that we implement the gestures not exactly as suggested but adopt certain of their characteristics (use of vertical hand position, work with fingers, pose or motion etc.) and define a new set of combinable gestures on this basis.

5. SUMMARY

Music production takes place in multi-device environments. Highly specialized hard- and software modules mold an often complex architecture. We discussed the role and integration of freehand gesture input in this scenario. Beyond the traditional interfaces that proved well for many tasks we regard freehand input a beneficial complement whenever it comes to continuous realtime control of multiple expressive parameters, e.g., for sound synthesis, audio effects and expressive performance.

As a first step toward the development of an appropriate set of gestures we conducted a survey with 44 participants. Besides the clear preference of one-handed gestures we collected several clues on which aspect of hand gestures (vertical hand movement, grab gesture and other finger movements, palm rotation) are favored for which type of musical parameter.

Acknowledgments We like to thank all visitors of the open house event who participated in the interview. This project is funded by the German Federal Ministry of Education and Research (BMBF 01UG1414A-C).

Bright/shrill (100% = 24 suggestions)

$S1$	spread fingers of one hand (4 suggestions, 16.7%)
$S2$	fast chaotic finger movements of one hand (2 suggestions, 8.3%)
$S3$	one open hand, palm heading downward, moves upward (2 suggestions, 8.3%)
$S4$	two claw hands, palms heading downward (2 suggestions, 8.3%)

Dark/soft (100% = 27 suggestions)

$s1$	smooth horizontal wave movement of one open hand, palm heading downward (3 suggestions, 11.1%)
$s2$	opposite of $S1$, one flat hand with closed fingers (2 suggestions, 7.4%)
$s3$	opposite of $S3$, one open hand, palm heading downward, moves downward (2 suggestions, 7.4%)
$s4$	both hands open, palms heading towards the computer, shaking, also described as repellent gesture to soften a shrill sound (2 suggestions, 7.4%)
$s5$	cover ears with both hands to soften the shrill sound (2 suggestions, 7.4%)
$s6$	swivel both open hands at the ears (2 suggestions, 7.4%)

Table 4. Suggested gestures for timbre control.

6. REFERENCES

[1] M. V. Mathews, "Three dimensional baton and gesture sensor," United States Patent Nr. 4,980,519, Dec. 1990.

[2] C. Havel and M. Desainte-Catherine, "Modeling an Air Percussion for Composition and Performance," in *NIME 2004*. Hamamatsu, Japan: Shizuoka University of Art and Culture, June 2001, pp. 31–34.

[3] L. Stockmann, A. Berndt, and N. Röber, "A Musical Instrument based on Interactive Sonification Techniques," in *Audio Mostly 2008*. Piteå, Sweden: Interactive Institute/Sonic Studio Piteå, Oct. 2008, pp. 72–79.

[4] Y. K. Lim and W. S. Yeo, "Smartphone-based Music Conducting," in *NIME 2014*. London, UK: Goldsmiths, University of London, June 2014, pp. 573–576.

[5] C. Kiefer, N. Collins, and G. Fitzpatrick, "Evaluating the Wiimote as a Musical Controller," in *ICMC 2008*. Belfast, Northern Ireland: ICMA, Sonic Arts Research Centre, Queen's University Belfast, 2008.

[6] N. Rasamimanana, F. Bevilacqua, N. Schnell, F. Guedy, E. Flety, C. Maestracci, B. Zamborlin, J.-L. Frechin, and U. Petrevski, "Modular Musical Objects Towards Embodied Control of Digital Music," in *TEI 2011*. Funchal, Portugal: ACM, 2011, pp. 9–12.

[7] K. Drossos, A. Floros, and K. Koukoudis, "Gestural User Interface for Audio Multitrack Real-time Stereo Mixing," in *Audio Mostly 2013*, Interactive Institute/Sonic Studio Piteå. Piteå, Sweden: ACM, Sept. 2013.

[8] T. Mitchell, S. Madgwick, and I. Heap, "Musical Interaction with Hand Posture and Orientation: A Toolkit of Gestural Control Mechanisms," in *NIME 2012*. Ann Arbor, Michigan, USA: University of Michigan, School of Music, Theatre & Dance, May 2012.

[9] S. Serafin, A. Trento, F. Grani, H. Perner-Wilson, S. Madgwick, and T. Mitchell, "Controlling Physically Based Virtual Musical Instruments Using The Gloves," in *NIME 2014*. London, UK: Goldsmiths, University of London, June 2014, pp. 521–524.

[10] S. Sridhar, "HandSonor: A Customizable Vision-based Control Interface for Musical Expression," in *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. Paris, France: ACM, Apr./May 2013, pp. 2755–2760.

[11] Uwyn bvba/sprl, "GECO: Multi-Dimensional MIDI/OSC/CopperLan Expression Through Hand Gestures," app on Airspace store, 2013, version 1.3.0.

[12] T. Murray-Browne and M. D. Plumbley, "Harmonic Motion: A Toolkit For Processing Gestural Data For Interactive Sound," in *NIME 2014*. London, UK: Goldsmiths, University of London, June 2014, pp. 213–216.

[13] M. M. Wanderley, "Gestural Control of Music," in *Proc. of the Int. Workshop on Human Supervision and Control in Engineering and Music*, Kassel, Germany, 2001, pp. 101–130.

[14] R. I. Godøy and M. Leman, Eds., *Musical Gestures: Sound, Movement, and Meaning*. New York, NY, USA: Routledge, Feb. 2010.

[15] E. R. Miranda and M. M. Wanderley, *New Digital Musical Instruments: Control And Interaction Beyond the Keyboard*, ser. The Computer Music and Digital Audio Series. Middleton, WI, USA: A-R Editions, Inc., July 2006, vol. 21.

[16] L. Dahl, "Triggering Sounds From Discrete Air Gestures: What Movement Feature Has the Best Timing?" in *NIME 2014*. London, UK: Goldsmiths, University of London, June 2014, pp. 201–206.

[17] J. Françoise, N. Schnell, and F. Bevilacqua, "A Multimodal Probabilistic Model for Gesture-based Control of Sound Synthesis," in *Proc. of the 21st ACM Int. Conf. on Multimedia*. Barcelona, Spain: ACM, 2013, pp. 705–708.

[18] J. Gossmann and M. Neupert, "Musical Interface to Audiovisual Corpora of Arbitrary Instruments," in *NIME 2014*. London, UK: Goldsmiths, University of London, June 2014, pp. 151–154.

[19] R. Carvalho and M. Neto, "Non Human Device #002," in *xCoAx 2014: Proc. of the 2nd Conf. on Computation, Communication, Aesthetics and X*, M. Carvalhais and M. Verdicchio, Eds., Faculdade de Belas Artes. Porto, Portugal: Universidade do Porto, June 2014, pp. 490–492.

[20] J. Han and N. Gold, "Lessons Learned in Exploring the Leap Motion™ Sensor for Gesture-based Instrument Design," in *NIME 2014*. London, UK: Goldsmiths, University of London, June 2014, pp. 371–374.

[21] L. Hantrakul and K. Kaczmarek, "Implementations of the Leap Motion in sound synthesis, effects modulation and assistive performance tools," in *Joint Conf. ICMC and SMC*, University of Athens. Athens, Greece: International Computer Music Association, Sept. 2014, pp. 648–653.

[22] A. Pon, J. Ichino, E. Sharlin, D. Eagle, N. d'Alessandro, and S. Carpendale, "VUZIK: A Painting Graphic Score Interface for Composing and Control of Sound Generation," in *ICMC 2012*, M. M. Marolt, M. Kaltenbrunner, and M. Ciglar, Eds. Ljubljana, Slovenia: International Computer Music Association, University of Ljubljana, Sept. 2012, pp. 579–583.

[23] D. Tormoen, F. Thalmann, and G. Mazzola, "The Composing Hand: Musical Creation with Leap Motion and the BigBang Rubette," in *NIME 2014*. London, UK: Goldsmiths, University of London, June 2014, pp. 207–212.

[24] J. Borchers, A. Hadjakos, and M. Mühlhäuser, "MICON a Music Stand for Interactive Conducting," in *NIME 2006*. Paris, France: IRCAM – Centre Pompidou, June 2006, pp. 254–259.

[25] M. T. Marshall, J. Malloch, and M. M. Wanderley, "Gesture Control of Sound Spatialization for Live Musical Performance," in *Gesture-Based Human-Computer Interaction and Simulation*, ser. Lecture Notes in Computer Science, M. Sales Dias, S. Gibet, M. M. Wanderley, and R. Bastos, Eds. Berlin, Heidelberg, Germany: Springer Verlag, 2009, vol. 5085, pp. 227–238.

[26] P. Quinn, C. Dodds, and D. Knox, "Use of Novel Controllers in Surround Sound Production," in *Audio Mostly 2009*. Glasgow, Scotland: Glasgow Caledonian University, Interactive Institute/Sonic Studio Piteå, Sept. 2009, pp. 45–47.

[27] W. Balin and J. Loviscach, "Gestures to Operate DAW Software," in *130th Audio Engineering Society Convention*. London, UK: Audio Engineering Society, May 2011.

[28] J. Ratcliffe, "Hand Motion-Controlled Audio Mixing Interface," in *NIME 2014*. London, UK: Goldsmiths, University of London, June 2014, pp. 136–139.

[29] M. Lech and B. Kostek, "Testing A Novel Gesture-Based Mixing Interface," *Journal of the Audio Engineering Society*, vol. 61, no. 5, pp. 301–313, May 2013.

[30] B. Owsinski, *The Mixing Engineer's Handbook*, 3rd ed. Boston, MA, USA: Course Technology, Cengage Learning, 2014.