A Review of Interactive Conducting Systems: 1970-2015

Kyungho Lee, Michael J. Junokas, Guy E. Garnett
Illinois Informatics Institute
University of Illinois at Urbana-Champaign
1205 W. Clark NCSA Building, Urbana, IL USA
klee141, junokas, garnett@illinois.edu

ABSTRACT

Inspired by the expressiveness of gestures used by conductors, research in designing interactive conducting systems has explored numerous techniques. The design of more natural, expressive, and intuitive interfaces for communicating with computers could benefit from such techniques. The growth of whole-body interaction systems using motion-capture sensors creates enormous incentives for better understanding this research. To that end, we retraced the history of interactive conducting systems that attempt to come to grips with interpreting and exploiting the full potential of expressivity in the movement of conductors and to apply that to a computer interface. We focused on 55 papers, published from 1970 to 2015, that form the core of this history. We examined each system using four categories: interface (hardware), gestures (features), computational methods, and output parameters. We then conducted a thematic analysis, discussing how insights have inspired researchers to design a better user experience, improving naturalness, expressiveness and intuitiveness in interfaces over four decades.

1. INTRODUCTION

In the history of Western art music, conductors have served as both physical and conceptual focal points. The modern form of conducting emerged due to the increasing complexity of symphonic scores over the nineteenth century. They became fully-fledged members of the performing ensemble, generating a stream of musical expression "running from composer to individual listener through the medium of the performer and further mediated by the expressive motions of the conductor." [1] In order to accomplish this goal, they used a variety of physical signatures to seamlessly convey musical expressions to the ensemble throughout rehearsals and performances. Conductors, in their increasingly complex task of directing the orchestra, have increasingly learned how to use embodied knowledge, as musicians and dancers did before them. Recent research supports this concept, showing that as a series of emblematic gestures, conducting has the capability of transmitting specific musical ideas, using a wide range of physical expressivity [2] [3].

With recent advances in sensing technology, the potential use of whole-body interaction (WBI) [4] plays a pivotal role in enhancing the natural user-interaction (NUI) paradigm, with an emphasis on embodiment. Since the field of WBI or NUI is relatively young and finds a novel interaction model to move researchers forward, conducting gestures have attracted researchers who seek fundamental insight into the design of complex, expressive, and multimodal interfaces. While the current natural-user interaction design paradigm has the ability to recognize the user's gestures and to operate a set of commands, it is still limited in extracting the expressive content from gesture, and even more limited in its ability to use this to drive an interactive system. The design of conducting interfaces has been driven by new methods or models that empower users through the augmentation of expression and/or expanding to a new degree of control to challenge the limitations. Our motivation is to start a systemic review of the history and state of the art, derived from these questions: What are the significant documents and experiments in the development of conducting systems? What is the research history and legacy of this domain? What can we learn from this body of research that might help us to design a better user experience?

Our paper will address interfaces that have been designed to capture conducting gestures, features and computational methods that have been used to interpret expressive contents in gestures, and strategies and techniques that have been used to define effective mappings from gesture to control of sound. Based on these points, we conducted a systematic review of fifty-five papers that used conducting gestures in interactive system design. This review is comprised of a sub-sample of papers related to interactive conducting systems that were selected from a broader literature search exploring the impact of designing multi-modal, expressive interfaces. A narrative review was additionally carried out in order to develop a more coherent understanding of expression-driven gesture design that supports human creativity, focusing on translating musical expression using the gesture. From this range of papers, three major themes in the history of designing interfaces with conducting gestures were addressed: naturalness, intuitiveness, and expressiveness. We described the keywords in the implications section in detail.

2. TERMINOLOGY

In this section, we present consensus-derived, fundamental concepts and definitions of interactive conducting systems, providing readers with a fundamental background for the better understanding the rest of the paper.

2.1 Interactive Conducting Systems

By referring to interactive 'conducting' systems, our focus narrows to a subset of interactive systems that use the breadth of standard, or typical, conducting gestures. Different researchers have defined the term in different ways. Early pioneers, for example, defined their systems as a conducting system [5], a music system, a conducting program [6], and a conductor follower [7]. In this paper, we define the term, *interactive conducting system*, as a system that is able to capture gestures from a conductor (as a user), extrapolate expressive contents in the gestures, assign appropriate meaning, and apply that meaning to the control of sound or other output media. Using such gestures, the conductor can manipulate a set of parameters interpreted by the system to produce outputs such as MIDI note/score playbacks, sound waveforms, and/or visual elements according to prescribed mapping strategies.

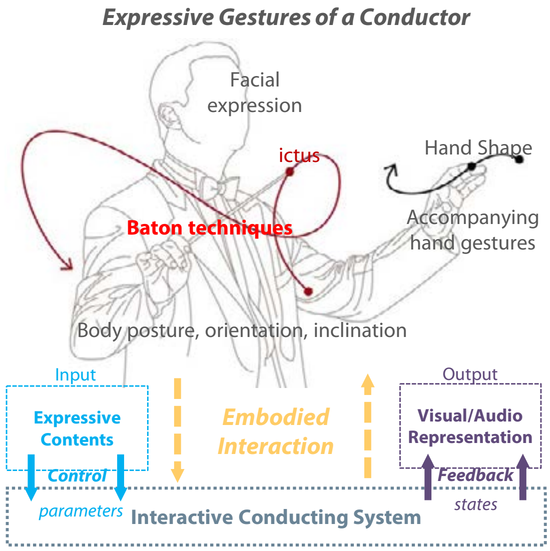

Figure 1. Illustration of an interactive conducting system, showing how conductors can drive a system using conducting gestures.

Figure 1 illustrates how an interactive conducting system works from our perspective. Note that the term 'embodied interaction' refers to using the perceivable, actionable, and bodily-experienced (embodied) knowledge of the user in the proximate environment (interactive system).

2.2 Conducting Gestures and Expressivity

A conductor uses expressive gestures to shape the musical parameters of a performance, interacting with the orchestra to realize the desired musical interpretation. While a conductor is directing, he or she makes use of diverse, often idiosyncratic, physical signatures such as facial expression, arm movement, body posture, and hand shape as seen in Figure 1. These physical signatures can convey different types of information simultaneously. Amongst these four different types of information channels, researchers have been mainly interested in the use of the hand and arm gestures in referring to conducting gestures, largely because these are the most standardized elements of the technique. Theoretically, conducting gestures have been investigated in linguistics as emblematic and pantomimic gestures according to the spectrum of the Kendon's continuum [8].

Based on this theoretical background, conducting gestures can be understood as a stream of linguistic information, which is relatively fixed, and lexicalized. There is very little variety in conveying specific musical direction to the others and decoding it from the gestures [9]. This view was addressed in Max Rudolf's authoritative conducting textbook [10] where he defined explicit parts of 'conducting gestures.' For example, conducting gestures can be classified into several groups by their intended effect (musical information) on the performance as done with baton techniques, which have been used to indicate the expression of each beat (e.g., legato, staccato, marcato, and tenuto), while accompanying left hand gestures have been used to support controlling dynamics, cues, cutoffs and vice versa.

There is a novel interpretation where the degree of variation in conducting gestures is used to enhance expression from the HCI perspective. Different conductors might perform the same musical expressions differently under the grammar of conducting. Therefore, we can consider expressivity is more associated with how to perform rather than what to perform.' Recent empirical research outcomes claimed that the individual variance or gestural differentiation can be understood as a degree of expression [2] providing a rich research area. Similarly, Caramiaux et al. [11] claimed that such differentiation can add "meaningful variation in the execution of a gesture" in expression-oriented interactions.

3. METHODOLOGY

3.1 Planning the Review

In this section, we identified the need for a systematic literature review, and developed a protocol that specifies methods to conduct data collection and analysis.

- Objective: Analyzing interactive conducting systems and computational methods used to find the challenges and opportunities for designing a better interactive systems, enabling the of use expressive, multimodal inputs.

- Research questions: (1) What types of interfaces have been designed to capture conducting gestures? (2) What features and computational methods have been applied to interpret expressive contents in gestures? (3) What strategies and techniques have been used to create effective mappings between these input gestures and applied outputs?

- Research sources: ACM, IEEE, CiteSeerX, SpringerLink, Computer Music Journal, Journal of New Music Research.

- Search strings: Our primary objective focuses on capturing and extrapolating expressivity from the conducting gestures, so we chose the following search strings after preliminary searches across the disciplines of musicology, psychology, machine learning, pattern recognition, and HCI studies. The first search string focuses on the design of interactive system that using conducting gestures. The second focuses on gesture recognition and analysis of conducting gestures using computational methods. The third focuses on the application of conducting gestures and movement.

1. conductor **and** (gesture or movement) **and** (interface **or** system) **or** (orchestra **or** ensemble) 2. conducting gesture **and** (expression **or** expressive gestures) **or** (recognition **or** analysis) 3. conducting gesture **and** (expression **or** expressive gestures) **or** (visual **or** sound)

- Language/Time restriction: Any papers published in English and available in the digital library.

- Inclusion criteria: (I1) Research comprising strategies, methods and techniques for capturing conducting gesture (conductor's gesture) and applying the results to design an interactive system/interface; (I2) Studies comprising theoretical backgrounds and computational methods to analyze and recognize characteristic aspects of conducting gestures; (I3) Projects using conducting gestures or conductors? expressions to drive interactive system to generate visuals/ audio.

- Exclusion criteria: (E1) Studies which do not meet any inclusion criteria; (E2) Studies focusing on conducting gestures using a computational approach but not related to any design aspect of HCI; (E3) Studies focusing on qualitative analysis of conducting gestures but not providing any computational methods; (E4) If two papers, from the same authors published in the same year, cover the same scope, the older one was excluded.

3.2 Conducting the Review

After defining a review protocol, we conducted the review. The data collection started at the beginning of 2015 with initial searches returning 129 studies with some overlapping results among the sources. After applying the inclusion, exclusion, and quality criteria, 55 papers were selected. The papers were primarily collected from ICMC (19 papers), ACM (4), IEEE (3), Computer Music Journal (2) and Journal of New Music Research (2). Other sources were collected from university data repositories (dissertation/thesis) or other journals.

4. RESULTS

Based on our investigation, we developed six different themes that have been centered around the history of interactive conducting systems: pioneers; tangible user interface; gesture recognition/machine learning; sound synthesis; commercial sensors; and visualization.

4.1 The first interactive conducting systems

Early interactive conducting systems design resorted to control and interaction paradigms of the time. They incorporated knobs, 3D joysticks, and keyboards as input devices. However, a series of pioneering explorations considering Engelbart's seminal demo [12] was presented in 1968. Mathews [13] described that his desire was to create an interface that would be able to connect the computer to the user as a conductor is connected to the orchestra. He fed the score information to the computer, which was paired with user interactions to make dynamic score interactions. Also, he adopted three modes (the score, rehearsal, and performance) that reflected the mental model of conductors. The name, a conducting system, was explicitly entitled by Buxton later in 1980. In Buxton et al.'s work [5], improved design considerations in terms of graphical representation were implemented. Such considerations enabled the user to adjust various musical parameters, such as tempo, articulation, amplitude, richness(timbre), on the screen through a textual user interface. The user controlled the parameters by typing numbers or moving cursors. These systems explored the potential of using non-conventional modalities and demonstrated how interactive conducting systems were being developed.

4.2 Rise of tangible user interface

Tangible interaction design generally encompasses user interfaces and interactions that emphasize "materiality of the interface; physical embodiment; whole-body interaction; the embedding of the interface and the users' interaction in real spaces and contexts." [14] Although this period was right before the explosion of tangible user interface design, we can see researchers' design reflecting its philosophy. From 1979, Mathews and Abbott [15] started designing a mechanical baton to use as an input device, allowing users to provide more intuitive input through its use. The baton was struck by the user with his or her hands or sticks and required no prior training for use. This tangible interface provided the user with the ability to capture the mental model of a conductor through the use of his or her embodied interaction with the machine. The consideration of tangibility and intuitiveness was advanced further by Keane et al. [16] starting in 1989. They designed a wired baton, which resembled an ordinary baton but was augmented with spring wires and an metal ball inside. By 1991, they improved the MIDI baton by adding a wireless transmitter and expanding the number of MIDI channels to 16, allowing the control of multiple parameters at the same time. Marrin et. al's Conductor's Jacket [17] further expanded this interface category. It is a wearable device that demonstrates the potential power of using EMG sensors, attempting to map expressive features to sections in the music score. Due to technological limits of this period, the overall weight of the device, including the digital baton, was a potential concern. In her later project, You're the Conductor [18] and Virtual Maestro [19], Marrin and her collaborators developed a gesture recognition system that was capable of mapping the velocity and the size of gestures to musical tempos and dynamics. Her approach has inspired numerous researchers interested in using the Arduino and accelerometers to measure the body's movement.

4.3 Use of Machine Learning Approach

In the history of interactive conducting systems, there have been three main challenges related with machine learning (ML): data collection, feature generation, and modeling. The first challenge was collecting conducting gestures and assuring the quality of this gestural dataset by removing outliers and smoothing signals. Many researchers needed to implement physical interfaces to measure the user's movement with higher precision. The second challenge was to find reliable and discriminative features to extrapolate expressivity from gestures including a dimensionality reduction process. A great deal of research has adopted kinematic features such as the velocity and acceleration to describe the movement. The third challenge was modeling the temporal dynamics of conducting gestures. Researchers have used Hidden Markov models (HMM) or neural networks (ANN) to create such models. Bien et al.'s work [20] was one of the first to adopt fuzzy logic to capture the trajectory of a baton in order to determine the beat. However, they built it based on the IF-THEN rule-based fuzzy system, not fully exploiting the potentials of fuzzy logic might have. Lee M. [7], Brecht [21] and their colleagues brought ANN to address conducting gesture recognition, using the Buchla Lightning baton [22] as an input device. They trained a two-layer multi perceptron (MLP) between six different marker points, time, and the probability of the next beat, using the ANN was adopted to deal with the local variations in conducting curves. Sawada et al. [23] [24] and Usa [25] also used ANN and HMM in their works respectively. In 2001, Garnett and his colleagues [26] advanced the algorithm by using distributed computing via open sound control(OSC), building on the success of the conductor follower. Kolesnik and Wanderley [27] proposed a system that captured conducting gestures using a pair of cameras, analyzing the images using EyesWeb. They used an HMM to recognize the beat and amplitude from the right and left hand expressive gestures. The exploration of ML approaches was accelerated by the advent of commercial sensors such as the Nintendo's Wiimote and the Microsoft's depth sensor, the Kinect V1 and V2. Bradshaw and Ng [28] adopted the Wiimote to analyze conducting gestures whereas other researchers [29] [30] [31] used the Kinect as an input sensor. Dansereau et al. [32] captured baton trajectories using a high quality motion capture devices (Vicon) and analyzed them by applying an extended Kalman filter as a smoothing method, using a particle filter for a training. Although the capability of capturing conducting gestures was advanced over time, the tracking results suggested that there were a lack of advancements in the input-output mappings, maintaining basic output parameters such as beat pattern, dynamics, and volume.

4.4 Sound Synthesis

One of the pioneering projects, GROOVE [13], was designed for "creating, storing, reproducing, and editing functions of time," for sound synthesis. After that, many researchers put their efforts into developing systems that enabled the user to control musical parameters in MIDI scores and audio files. Their projects allowed users to directly manipulate musical performances, mapping kinetic movements to sound. Morita et al. [33] began realizing a system that gave "an improvisational performance in real-time." To achieve their goal, they adopted computer vision technology to track the conductor's baton. With the system, the user can manipulate tempo, strength (velocity), start, and stop of the music. In following work, they extended the system, adding a data glove to capture additional expressions of hand shapes. From 2001, Borchers et al. [34] presented a series of "personal orchestra" projects, allowing the user to control tempo, dynamics, and instrument emphasis based on pre-recorded audio files. During the same period, Murphy et al. [35] and Kolesnik [27] attempted to implement systems to play time-stretched sound in real time using a variant of the phase vocoder algorithm. However, computing power was not sufficient to guarantee synchronous audio and video playback, so video or audio playback module were dealt with independently. In this regard, Lee and his colleagues' work had significantly contributed to addressing these problems. He described his concept as semantic time [36], aiming to allow the user to perform time-stretching without substantially losing or distorting the original information. He applied the technique to multiple projects: *conga*, *You're conductor* and *iSymphony* [37] [18]

4.5 Advent of commercial sensors

Until the 2000's, many researchers investigated a conductor's gestures by attaching customized sensors to body parts or analyzing motion in a lab context to acquire the highest quality of datasets. However, the advent of relatively cheap and robust sensors, such as Nintendo Wiimote and Microsoft Kinect, led researchers to a different approach. Nintendo introduced the Wiimote in late 2006 as an advanced input device incorporating a 3-axes accelerometer and infrared sensor. It supported the Bluetooth protocol for communication. Microsoft Kinect, which was presented in 2009 for V1 and 2014 for V2, was featured an RGB camera, depth sensor, and a microphone array. One of the primary reason for adopting commercial sensors is that they are less expensive, non-invasive, yet powerful and can be used in general contexts which accelerates the data collection and iterative design process. By the year of 2000, many research projects had been designed to use these sensors. Bradshaw and Ng [28] used multiple Wiimotes to capture 3D acceleration data of conducting gestures. They attempted to extract information and use the parameters to change tempo and dynamic then feed them back to the user using several appropriate methods including sonification, visualization and haptics (i.e vibration in the controller). Toh et al. [29] designed an interactive conducting system using the Kinect V1, allowing the user to control tempo, volume, and instrument emphasis. It was also one of the first attempts at using a body posture for control information. Rosa et al. [30] designed another system that allowed the user to conduct a virtual orchestra, controlling the tempo, the overall dynamics, and the specific volume levels for sets of instruments in the orchestra.

4.6 Visualization of expressivity

Unlike the other advancements in the history of designing interactive conducting systems, little attention has been paid to visualizing the dynamics of conducting gestures and its expressivity. The uncharted territory is challenging due to: 1) the concrete conceptual model that lead researchers to understand the qualitative aspect of conducting gestures. 2) the feature generation and recognition methods to analyze and extract expressivity from the movements. Nevertheless, there were several attempts to visualize some dimensions of conducting gestures. One of the early attempts was made in Garnett et al.'s project [38], Virtual Conducting Practice Environment. They visualized the four beats in 4/4 beat pattern and the horizontal line representing the beat plane. In 2000, Segen and Gluckman [39] presented their project, Visual Interface for Conducting Virtual Orchestra, at SIGGRAPH. While the MIDI sequencer was playing an orchestral score, the user was able to adjust its tempo and volume. 3D human models were rendered and animated, that follow pre-designed movements and choreography, based on the tempo set. Bos et al. [40] implemented the virtual conductor system that conducted music specified by a MIDI file to human performers. It received input from a microphone, responding to the tempo of the musicians. This was the first use of a virtual agent to direct other human agents instead of being controlled by the user. Recently, Lee et al. [41] created an interactive visualization to represent expressivity of the conducting gestures. They adopted Laban Movement Analysis to parameterize expressivity. The visualization received an input video stream and was driven by expressive motion parameters extracted from the user gestures, rendering particle graphics.

5. IMPLICATIONS

Based on the synthesis of the survey, we drew three implications for future design works. These implications reflect the current trend of designing WBI/NUI paradigm based on Norman and van Dam's note. Norman proposed that designers could improve user performance by mapping knowledge in the world to expected knowledge in the user's mind [42]. van Dam suggested that the ideal user interface "would let us perform our tasks without being aware of the interface as the intermediary." [43] Upon consideration, the future of interactive conducting systems should consider the three core elements one step further: 1) *naturalness* which is allowing a multi-limbed and multimodal interaction; 2) *intuitiveness* which is enabling embodied interaction; 3) *expressiveness* which is inspiring the user's creative tasks through transmodal feedbacks. We describe each implication in more detail.

5.1 For Being Natural

Amongst several definitions, we can define being natural in our context as a sensing technique for having more holistic forms of inputs that allow the user to use multi-limbed and multi-modal interaction. With advanced sensing mechanisms, we witnessed that new forms of 'natural' input have arisen to replace traditional WIMP based mechanisms. With machine learning techniques, the whole-body interaction can make the best use of our embodied abilities and real world knowledge [44]. However, our analysis results suggest that we need to explore other techniques to extrapolate expressivity in conducting gestures revealed not only through movement, but also through facial expressions, muscles tensions, or brain activities. Because current models and sensors are not sensitive enough to extrapolate affective or cognitive states from subtle gestures (external cues) that represent the internal cognitive or affective state indirectly [45]. In addition to sensing external cues, we can consider adopting Brain-Computer Interfaces (BCI) to capture significant insight from the users' emotional state more directly. By adopting BCIs, we can utilize rich information not only to operate a set of commands with the user's brain activity instead of using motor movements but also to provide more natural ways of controlling interfaces. For example, recalling a pleasant moment could be recognized and interpreted as expression parameters to control the system in the highest possible natural and intuitive manner.

5.2 For Being Intuitive

Raskin [46] argued that an 'intuitive' interactive system should work in a similar way that the user does without pre-training or rational thought. He suggested that a user interface could incorporate intuitiveness by designing towards (even identically) something the user already knows. In the history of the interactive conducting systems, numerous researchers have designed tangible interfaces and created visualizations that resembled the real-world context of conductors to keep their mental model as similar as possible developed under the term of intuitive design. We propose to put more consideration on embodied interaction in the design process. A growing body of research in the understanding body-mind linkages has supported this claim, explaining how abstract concepts and ideas can become closely tied to the bodily experiences of sensations and movements. In the HCI fields, Höök [47] provided evidence of how "our corporeal bodies in interaction can create strong affective experiences." It is expected that the embodied interaction design approach will improve the overall user experience and the performance of conducting machines. As Norman [48] noted, designers can improve user performance with the interactive system by providing a better mapping knowledge from the world (determined by system design) to expected knowledge in the user's head.

5.3 For Being Expressive

Dobrian claims [49], musical instruments or interfaces cannot be expressive as they do not have anything to express until the user commands what to express and how to express it. However, we observed a great deal of ideas utilizing computers as a vehicle to transmit a conductor's expressiveness to the machine and to the audience in the history. Researchers have explored a variety of ways to quantify conductors' gesture and to transform the significance of expressivity into a mental musical representation. The exploration can be interpreted as a journey of designing creativity support tools in the music domain as we saw with many researchers experimenting with scores composed in a MIDI or waveforms producing the different quality of sound in their evaluation process. Our analysis demonstrated that only very few visual explorations were made through the history of interactive conducting, and further exploration is rich with opportunity. In this context, the concept of metacognition gives us evidence to consider adoption since it explains how our cognitive system evaluates and monitors our own thinking processes and knowledge content [50]. Research findings showed that the metacognitive feeling of knowing, so-called confidence, can help the users to associate possible ideas together, guiding the users to a path to accomplish the goal [51].

6. CONCLUSION AND FUTURE WORK

We found that numerous interactive conducting systems had been researched and implemented over forty years reflecting the emerging technologies and paradigms from the HCI. Interactive conducting systems explore numerous, different approaches to making the best use of expressivity in conducting gestures from different perspectives; the kinematics of conducting gestures associated with tracking beats; the recognition of particular types of conducting gestures including articulation styles; and mapping for music control or synthesis. The interactive conducting systems were also developed and evaluated for various purposes such as performance, pedagogy, and scientific research prototypes to validate theory or algorithms. With three design implications, we can imagine the possible interactive system scenarios such as: 1) 'a machine symphony' which enables the conductors (the users) to lead a full-size orchestra made of 70-100 high quality virtual instruments based on MIDI scores; 2) 'an augmented ensemble' which visualizes expressivity in the conductors' movement through augmented/mixed reality technology; and 3) 'a pedagogical agent' that helps the users' embodied learning process for basic components of conducting gestures such as beat patterns and articulations styles.

Acknowledgments

This research was supported by the Social Sciences and Humanities Research Council of Canada (SSHRC).

7. REFERENCES

[1] C. Small, "Musicking–the meanings of performing and listening," vol. 1, no. 1, p. 9.

[2] G. Luck, P. Toiviainen, and M. R. Thompson, "Perception of expression in conductors' gestures: A continuous response study."

[3] G. D. Sousa, "Musical conducting emblems: An investigation of the use of specific conducting gestures by instrumental conductors and their interpretation by instrumental performers," Ph.D. dissertation, The Ohio State University, 1988.

[4] D. England, M. Randles, P. Fergus, and A. Taleb-Bendiab, "Towards an advanced framework for whole body interaction," in *Virtual and Mixed Reality*. Springer, pp. 32–40.

[5] W. Buxton, W. Reeves, G. Fedorkow, K. C. Smith, and R. Baecker, "A microcomputer-based conducting system," pp. 8–21.

[6] R. B. Dannenberg and K. Bookstein, "Practical Aspects of a midi conducting program," in *Proceedings of the 1991 International Computer Music Conference*, pp. 537–40.

[7] M. Lee, G. Garnett, and D. Wessel, "An adaptive conductor follower," in *Proceedings of the International Computer Music Conference*. International Computer Music Association, pp. 454–454.

[8] D. McNeill, *Gesture and thought*. University of Chicago Press.

[9] A. Gritten and E. King, Eds., *New perspectives on music and gesture*, ser. SEMPRE studies in the psychology of music. Farnham ; Burlington, VT: Ashgate Pub, 2011.

[10] M. Rudolf, *The grammar of conducting: a practical guide to baton technique and orchestral interpretation*. Schirmer Books.

[11] B. Caramiaux, M. Donnarumma, and A. Tanaka, "Understanding Gesture Expressivity through Muscle Sensing," vol. 21, no. 6, p. 31.

[12] D. C. Engelbart and W. K. English, "A research center for augmenting human intellect," in *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. ACM, 1968, pp. 395–410.

[13] M. V. Mathews and F. R. Moore, "GROOVE—a program to compose, store, and edit functions of time," vol. 13, no. 12, pp. 715–721.

[14] B. Ullmer and H. Ishii, "Emerging frameworks for tangible user interfaces," *IBM systems journal*, vol. 39, no. 3.4, pp. 915–931, 2000.

[15] M. V. Mathews and C. Abbott, "The sequential drum," pp. 45–59.

[16] D. Keane, *The MIDI baton*. Ann Arbor, MI: MPublishing, University of Michigan Library.

[17] T. Marrin and R. Picard, "The 'Conductor's Jacket': A Device for Recording Expressive Musical Gestures," in *Proceedings of the International Computer Music Conference*. Citeseer, 1998, pp. 215–219.

[18] E. Lee, T. M. Nakra, and J. Borchers, "You'Re the Conductor: A Realistic Interactive Conducting System for Children," in *Proceedings of the 2004 Conference on New Interfaces for Musical Expression*, ser. NIME '04. Singapore, Singapore: National University of Singapore, 2004, pp. 68–73.

[19] T. M. Nakra, Y. Ivanov, P. Smaragdis, and C. Ault, "The ubs virtual maestro: An interactive conducting system," pp. 250–255.

[20] Z. Bien and J.-S. Kim, "On-line analysis of music conductor's two-dimensional motion," in , *IEEE International Conference on Fuzzy Systems, 1992*, pp. 1047–1053.

[21] B. Brecht and G. Garnett, "Conductor Follower," in *ICMC Proceedings*, 1995, pp. 185–186.

[22] R. Rich, "Buchla Lightning MIDI Controller: A Powerful New MIDI Controller is Nothing to Shake a Stick at," *Electron. Music.*, vol. 7, no. 10, pp. 102–108, Oct. 1991.

[23] H. Sawada, S. Ohkura, and S. Hashimoto, "Gesture analysis using 3D acceleration sensor for music control," in *Proc. Int'l Computer Music Conf.(ICMC 95).*

[24] T. Ilmonen and T. Takala, *Conductor Following With Artificial Neural Networks*.

[25] S. Usa and Y. Mochida, "A conducting recognition system on the model of musicians' process," vol. 19, no. 4, pp. 275–287.

[26] G. E. Garnett, M. Jonnalagadda, I. Elezovic, T. Johnson, and K. Small, "Technological advances for conducting a virtual ensemble," in *International Computer Music Conference,(Habana, Cuba, 2001)*, pp. 167–169.

[27] P. Kolesnik and M. Wanderley, "Recognition, analysis and performance with expressive conducting gestures," in *Proceedings of the International Computer Music Conference*, pp. 572–575.

[28] D. Bradshaw and K. Ng, "Analyzing a conductors gestures with the Wiimote," pp. 22–24.

[29] L. Toh, W. Chao, and Y.-S. Chen, "An interactive conducting system using Kinect," 2013, dOI: 10.1109/ICME.2013.6607481.

[30] A. Rosa-Pujazon, I. Barbancho, L. J. Tardon, and A. M. Barbancho, "Conducting a virtual ensemble with a kinect device," in *Proceedings of the Sound and Music Computing Conference 2013*, ser. SMC'13. Logos Verlag Berlin, pp. 284–291.

[31] Á. Sarasúa and E. Guaus, "Dynamics in music conducting: A computational comparative study among subjects," in *14th International conference on New interfaces for musical expression*, ser. NIME'14, vol. 14.

[32] D. G. Dansereau, N. Brock, and J. R. Cooperstock, "Predicting an orchestral conductor's baton movements using machine learning," vol. 37, no. 2, pp. 28–45.

[33] H. Morita, S. Hashimoto, and S. Ohteru, "A computer music system that follows a human conductor," vol. 24, no. 7, pp. 44–53.

[34] J. O. Borchers, W. Samminger, and M. Mühlhäuser, "Conducting a realistic electronic orchestra," in *Proceedings of the 14th annual ACM symposium on User interface software and technology*. ACM, pp. 161–162.

[35] D. Murphy, T. H. Andersen, and K. Jensen, "Conducting audio files via computer vision," in *Gesture-based communication in human-computer interaction*. Springer, pp. 529–540.

[36] E. Lee, T. Karrer, and J. Borchers, "Toward a framework for interactive systems to conduct digital audio and video streams," *Computer Music Journal*, vol. 30, no. 1, pp. 21–36, 2006.

[37] E. Lee, I. Grüll, H. Kiel, and J. Borchers, "conga: A framework for adaptive conducting gesture analysis," in *Proceedings of the 2006 conference on New interfaces for musical expression*. IRCAM—Centre Pompidou, pp. 260–265.

[38] G. E. Garnett, F. Malvar-Ruiz, and F. Stoltzfus, "Virtual conducting practice environment," in *Proceedings of the International Computer Music Conference*, pp. 371–374.

[39] J. Segen, J. Gluckman, and S. Kumar, "Visual interface for conducting virtual orchestra," in *15th International Conference on Pattern Recognition, 2000. Proceedings*, vol. 1, pp. 276–279 vol.1.

[40] P. Bos, D. Reidsma, Z. Ruttkay, and A. Nijholt, "Interacting with a Virtual Conductor," in *Entertainment Computing - ICEC 2006*, ser. Lecture Notes in Computer Science, R. Harper, M. Rauterberg, and M. Combetto, Eds. Springer Berlin Heidelberg, no. 4161, pp. 25–30.

[41] K. Lee, D. J. Cox, G. E. Garnett, and M. J. Junokas, "Express It!: An Interactive System for Visualizing Expressiveness of Conductor's Gestures," in *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, ser. C&C '15. New York, NY, USA: ACM, 2015, pp. 141–150.

[42] D. A. Norman, *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[43] A. van Dam, "Beyond WIMP," *IEEE Computer Graphics and Applications*, vol. 20, no. 1, pp. 50–51, 2000.

[44] D. England, "Whole Body Interaction: An Introduction," in *Whole Body Interaction*, ser. Human-Computer Interaction Series, D. England, Ed. Springer London, pp. 1–5.

[45] D. Tan and A. Nijholt, "Brain-Computer Interfaces and Human-Computer Interaction," in *Brain-Computer Interfaces*, ser. Human-Computer Interaction Series, D. S. Tan and A. Nijholt, Eds. Springer London, pp. 3–19, DOI: 10.1007/978-1-84996-272-8_1.

[46] J. Raskin, "Intuitive equals familiar," vol. 37, no. 9, pp. 17+.

[47] K. Höök, "Affective loop experiences: designing for interactional embodiment," vol. 364, no. 1535, pp. 3585–3595.

[48] D. A. Norman, *The design of everyday things*, 1st ed. Doubleday,.

[49] C. Dobrian and D. Koppelman, "The'E'in NIME: musical expression with new computer interfaces," in *Proceedings of the 2006 conference on New interfaces for musical expression*. IRCAM—Centre Pompidou, pp. 277–282.

[50] C. Hertzog and D. F. Hultsch, "Metacognition in adulthood and old age."

[51] T. Bastick, *Intuition : how we think and act*. J. Wiley,.