Revista Vórtex | Vortex Music Journal | ISSN 2317–9937 | http://vortex.unespar.edu.br/
DOI: https://doi.org/10.33871/vortex.2025.13.9841

ORIGINAL PAPER

Music Conducting Pedagogy and Technology: A Descriptive Review of Literature

Adamilson Guimarães de Abreu

Universidade Federal do Pará, Escola de Música | Belém, Pará, Brazil

Resumo: Esta revisão da literatura descreve a relação entre o desenvolvimento de tecnologias e o ensino da regência musical, considerando que gestos musicais são um dos importantes aspectos da regência já que informam a forma, a articulação e as intenções de expressividade da música. Artigos acadêmicos sobre regência e o uso de tecnologias tais como aplicativos de celular, redes de internet, robôs e inteligência artificial (IA) foram selecionados com objetivo de elucidar quais ferramentas tecnológicas estão em desenvolvimento e como podem estar sendo assimiladas às práticas pedagógicas. Como método de análise foi utilizado o enfoque da comparação constante, incluindo a síntese e a interpretação dos artigos. Resultados sugerem que desenvolvedores de tecnologias continuam a explorar modelos no sentido de substituir parcialmente a interação direta professor-aluno no ensino da regência musical. No entanto, as diversas possibilidades exploradas por tecnologistas estão à espera de se tornarem ferramentas tecnológicas disponíveis em contextos de ensino.

Palavras-chave: Regência musical, Pedagogia da regência musical, Reconhecimento de gestos musicais, Tecnologia.

Abstract: This literature review intends to describe the relationship between technological developments and the teaching of music conducting techniques considering that musical gestures are one of the most important aspects since they inform the music's form, articulation, and overall expressiveness of the music. Research articles were purposefully selected to collect information about how current technologies such as smartphone applications, virtual reality, robots, and artificial intelligence (AI) are being or not incorporated into the teaching of music conducting. I used the constant comparative approach to perform this review, including synthesizing and interpreting the selected research articles. Results suggest that technology developers continue to explore models to substitute human-to-human interaction regarding the teaching of music conducting partially. Nevertheless, the diverse possibilities explored are still waiting to become actual tools available in music-conducting teaching settings.

Keywords: Music conducting, Music conducting pedagogy, Musical gestures recognition, Technology.

Music Conducting Pedagogy and Technology: A Descriptive Review of Literature. Recebido em 29/10/2024. Aprovado em 29/04/2025. Disponível online: 09/10/2025. Editor: Dr. Felipe de Almeida Ribeiro; Dr. Fabio Guilherme Poletto.

Conductors communicate ideas through physical gestures and movements. The brain function that detects bodily positions or movements of muscles is termed kinesthesia (Proske & Gandevia, 2012). Conducting gestures have developed in various ways as a kinesthetic music practice over time. Southerland (2019) wrote about how conducting gestures evolved historically over the years and noted that "depictions of conducting gestures were etched in hieroglyphics and Pharaonic friezes, recorded in Greek and Roman literature and encoded above Jewish and Gregorian texts as cheironomic neumes" (p. 30).

Lately, technological advances have entered the realm of music conducting. Music Conducting resembles communication models that communicate artistic concepts and technical parameters. In that regard, the use of technology models such as Virtual Reality Learning Environment (VRLE) and tools such as smartphone applications are challenging conducting practice, pedagogy, and research (Ferguson & Wanderley, 2010; Fazekas et al. (2013). Unfortunately, the knowledge and information experts possess are not meeting the growing demands of aspiring music technologists (Brader, 2009). Furthermore, while technology constantly advances, the pedagogy of music conducting may not be in synchrony with such technological developments as far as teaching practices. This literature review is an effort to understand better such contemporary phenomenon and to provide conducting teachers and music educators with information that might aid in the challenge of applying technological developments to classroom activities regarding music conducting.

1. Method

This descriptive literature review is part of the research materials featured in my doctoral dissertation. I chose to apply the method of document analysis using the constant comparative approach (Bowen, 2009) to this literature review because documents in general, and research articles in particular, help provide data if analytical procedures are applied. As Merriam (1988) wrote, "Documents of all types can help the researcher uncover meaning, develop understanding, and discover insights relevant to the research problem" (p. 118).

1.1. Data Sources and Sampling

According to Glaser and Strauss (1967), the selection of publication material is the equivalent of a collection of interviews or field notes. The gathering and selection of articles lasted from January of 2021 to May of 2022. During this period, I searched publicly available resources such as *Google Scholar* in addition to peer-reviewed journals such as *Frontiers in Digital Humanities* and *Journal of Music, Technology, and Education*. Other sources included paper presentations at conferences and symposiums about music and technology such as the *International Conference on New Interfaces for Musical Expression (NIME)*, the *International Symposium on Computer Music Modeling and Retrieval,* and the *Institute of Engineering, Electrotechnicians and Electronics (IEEE) International Conferences*. I searched these databases using the following keywords: conducting, music conducting, music conducting pedagogy, musical gestures, gesture recognition, and technology, and used each keyword alone and in combination with one another.

To discover additional resources, I performed a manual search in the reference list of relevant articles that I encountered using the same keywords. I also chose materials that matched one of the following criteria: (a) the articles were perceived to be of pedagogical interest for conducting music teacher educators or (b) the authors of the articles exploring the intersection between conducting gestures and the use of technology. These materials were chosen so that I could gain insights into the application of technology models such as digital and robotic orchestras-conductors, and gestural controlling-analytical systems in music conducting research context because they might inform conducting instructors' pedagogical practices. All the collected resources were in the English language.

I adopted a two-stage process as a strategy to select the included articles. First, only the abstract of the papers was studied to make the first selection. Second, the full text was screened to ensure that each one of the research articles matched the previously reported selection criteria (Snyder, 2019). Although matters involving the surge of new technologies are an ever-developing subject, the art and craft of music conducting is not. Music conducting is a traditional art form in which technique and pedagogy have been slowly consolidated since the end of the nineteenth century. Therefore, to find a balance between the traditional use of music conducting techniques and the necessity of being UpToDate in matters of technological developments, my search included articles from 1990-2022.

Once the subjects as well as the authors of the articles started to repeat themselves according to the pre-stablished selection criteria and timeframe, the search for new articles stopped. The timeframe utilized to review the literature did not apply to the articles that support the research methodology. Thus, the selected research articles aggregated during a 32-year time lapse have the potential to soundly represent a contemporary voice of music conducting researchers about the teaching and learning of conducting techniques and musical gestures regarding the use teaching assistive technologies.

1.2. The Analysis Procedures

I chose to use the method of document analysis to ground my literature review in the context of music-conducting pedagogy and technological developments and trends. Document analysis is a process of evaluating documents to produce empirical knowledge and develop an understanding of a phenomenon (Bowen, 2009; Corbin & Strauss, 2008). This form of research uses documents that are interpreted by the researcher to give voice and meaning to a particular topic of interest. Analyzing documents includes coding content into themes similar to how focus groups or interview transcripts are analyzed (Bowen, 2009).

Document analysis involves written forms of communication (Kuckartz, 2014) and focuses on what is contained within these documents (e.g., text, figures, and/or tables) (Benaquisto & Given, 2008). This process encompasses the analysis of the content of a wide range of documents, such as texts, official publications and reports, letters, and written responses to open-ended surveys, including research articles (Patton, 2002). Content is generally not under the investigators' control when performing a document analysis. Therefore, data may need to emerge from the material itself rather than imposed a priori by a theoretical construct.

In sum, document analysis enables researchers to investigate human behaviors indirectly by analyzing products such as textbooks, articles, newspapers, advertisements, songs, political statements, novels, and pictures—virtually all types of communication (Aktürk & Demircan, 2017). As a technique for the analysis, I used content analysis (Braun & Clarke, 2006) in combination with the constant comparative approach (Bowen, 2009) because it enabled identifying, analyzing, and reporting patterns in the form of themes within and amongst texts. From that perspective, the immerged categories of this literature review were grounded in the data within its specific contexts as follows.

2. Human and Machine Learning

With the development of artificial intelligence and robotics, our psychological and sociological understanding of what it is that makes us distinctive and unique has evolved. For instance, body language, including facial expression, is an embodied skill that is difficult to describe, but easy to demonstrate. Humans learn by observing other humans performing tasks. Learning action sequences or operation plans—such as when a child observes an adult getting a bottle from the refrigerator and repeats the learned action––is an abstract problem, which supposes modeling of cognitive skills (Dillmann, 2004).

In this sense, there is an ongoing paradigm shift about the way machines perform tasks. Ansari et al. (2018) encourage us to rethink the human-machine because this process affects both human workers and intelligent machines. How does it work? "From a cognitive computer perspective, artificial models and computational algorithms resemble the ability of human learning and reproduce human skills" (p. 118).

Human-centered ML is programming a computer to perform a task by providing human examples rather than by writing a program code describing the behavior (Gillies et al., 2016). For instance, instead of describing computational procedures that reason about what features make something a human face and translating that idea into program codes, human-centered ML works on face recognition systems out of a large database of face images. While using such a database, ML technology can create virtual characters who use appropriate facial expressions to interact with humans by recognizing categories of human actions sensed with sensing mechanisms or create expressive gestural controllers with a particular feeling (Fiebrink, 2011, as cited in Gillies et al., 2016).

Considering this ML paradigm, we should reflect upon how we think about human learning. Dillman (2004) suggested that humans (a) navigate through changing environments, (b) adapt recognition abilities to a particular scenery, (c) manipulate a wide range of objects, and (d) teach themselves what to do and how to do it by acquiring relevant information through observation and multimodal dialogs. In addition, humans exchange learning-teaching experiences and distinguish parameters for a particular demonstration specific to the problem's concept. Finally, we transfer observed information from a specific task to an independent and flexible knowledge structure that can be generalized to new situations. In fact, robots are rapidly developing to mimic such cognitive processes.

Ansari et al. (2018) argued that the only aspect that distinguishes the capability of humans and machines in performing an assigned task is the subjective quality and performance variations in carrying out such a task. With interactive ML, the user chooses what new examples to label for the algorithm or works together with the computer in controlling the process. The computer is part of the human design process to improve the robot's tasks (Gillies et al., 2016). Importantly, by training algorithms to use high-level human-machine interfaces, users can build and refine systems for real-time gestural control. Although we are refining our understanding of how we learn by teaching machines to learn just the way humans do, the ML concept is already being applied within music conducting and technology research.

Before I review the literature on music conducting and technology in more detail, two concepts need to be explored, as they underpin the body of research on music conducting pedagogy and technology, namely the concept of Music Affordance (MA) and Network Music Performance (NMP).

2.1. Music Affordance

Affordance was originally conceived to describe relationships between the environment and an actor–a person who interacts with the environment. According to Tanaka and Altavilla (2012), music affordance (MA) is a configuration of properties that provide a direct link between perception and action related to technological music interfaces. In music, affordance can be configured as screen-based interfaces functioning as music digital instruments because sounds carry information for affordance that cannot be seen (Gaver, 1991).

Gestural affordance has complex, multi-faceted influences like size, form, culture, history, and personal experience. Godøy (2009) explained the theoretical concept of *sound tracing* behind this idea: the sound of an instrument affords gestures that imitate actual musical practice. People imagine the afforded object to be attached to the instrument the sounds refer to. Conversely, unfamiliar objects allow for exploration of what can be done with them as far as producing sounds. Tanaka and Altavilla (2012) also described that a specific musical process, such as music conducting, demands engagement with cognitive and cultural processes which means affordance is also shaped by a person's previous knowledge and experience. In other words, the context of the use affects how the affordance of an object is perceived by the actor. Thus, music technologists who explore technology tools to teach conducting likely consider the MA concept in their experiments and tests.

For example, Tanaka and Altavilla (2012) performed a comparative study of gestural interaction with a musical sound. They conducted an interview-based user study involving three accelerometer-based devices: an *iPhone*, a *Wii-remote,* and an *Axivity Wax* prototype–a small, wireless hand device with the size of the head of a baton, with musical sounds including percussion, stringed instruments, and voice recordings. By using computer mappings across different source sounds, and performing them from three different devices, users experienced forms of physical, sonic, and cultural affordance that combined to form MA. Findings were that the least familiar one out of the three accelerometers, the *Axivity Wax* prototype, was perceived as giving more freedom to participants, allowing them to focus on sound production which was disassociated with the object's shape and form. Thus, the researchers concluded that the form and shape of the objects informed MA. Given the brain's capability of mentally rehearsing movements and gestures, MA is a very

 | https://doi.org/10.33871/vortex.2025.13.9841

important concept regarding human-computer interaction. Understanding brain functions when certain aspects of music conducting practice such as arm/hand gestures, torso movement, and eye contact are to be performed might illuminate the teaching of conducting using e-learning processes and devices such as accelerometers.

Just as human-computer interaction has developed and affected the concept of MA, human-to-human interaction has also been affected by technology. One prominent example includes the digital interaction among music performers on the internet, with implications for music conducting pedagogy and technology.

2.2. Network Music Performance

Network Music Performance (NMP) is a concept about the use of internet applications for music interaction that aims to enable musicians to collaborate remotely through a telecommunication network (Rottondi et al., 2016). NMP is sound without space. In an ideal NMP situation, musicians and audience members are placed in a virtual environment to experience musical events by collaborating and communicating from different locations, all interconnected by high-fidelity multichannel networked systems (Sawchuk, 2003). The first goal of NMP is to have musicians in multiple locations around the world performing together in real-time using high-speed internet, with no latency, in front of a live audience (Kapur et al., 2005). Technology developers in the NMP field want to create a complete aural and visual environment that places people in a virtual space where they can communicate naturally from different physical locations, but where there are no technological difficulties.

2.2.1. Challenges for Implementing NMP Systems

Network transmission presents several limitations to virtual live music performances. According to Zimmermann et al. (2008), audio, sound, and potentially haptic data–any technology that can create an experience of touch or any form of technological interaction involving touch (Rodet et al., 2005) –are considered "isochronous media streams" (p. 142), or data that cannot be stored and needs to be transmitted in a critical time fashion for live streaming. That reality imposes a series of technical challenges for NMP such as (a) network delay during audio and video acquisition and rendering due to dropped or delayed packets, (b) contention between network interfaces such as sound cards and CPU units that retain data during network communication, and (c) two-way transmission that causes echo cancelations over the network. These technical events impose signal synchronization strains on the system. They interrupt the natural visual and aural feedback that musicians exchange with each other when performing to guarantee a synchronized execution of the music (Holub et al., 2012; Kapur et al., 2005; Sawchuk et al., 2003; Zimmermann et al., 2008).

Latency is one of the biggest challenges concerning NMP. Musicians tolerate latency and jitter below a few tens of milliseconds, estimated between 20-30ms (compared to 8-9 m distance between musicians on stage––the maximum separation they can synchronize without a conductor) (Rottondi et al., 2016). Chamber orchestras can only tolerate 10-40 ms of latency. Jazz musicians can work around 80 ms and still be able to synchronize in concerts, whereas general-purpose communication tolerates up to 150 ms of delay (Sawchuk et al., 2003). Latency tolerance is also correlated with the type of instrument and tempo of the music. For example, organists are used to a 20-30ms delay between pressing the note on the keyboard and the actual sound coming out of the pipes. Slower tempi are three times more tolerable (150ms delay) than fast music (50ms). In between these extremes, it seems there is a window of 50-100 ms "segmental tempo range" where musicians attempt to synchronize despite latency (Zimmermann et al., 2008, p. 14). Hearing their instrument delayed causes more discomfort than hearing another musician playing late because they also use eye contact and body language–torso, arm, and head movements–to signal. Because musicians are used to body language to deal with sound delay, conductors are an essential part of the synchronizing process during performance, especially in large ensembles. Conductors become the focal point where all the musicians look aiming to play together because synchronization is not limited by the sound they hear but guided for the visual orientation of the anticipated conducting gestures.

2.2.2. Coping with NMP Limitations

Because technological strains in NMP are currently unavoidable, musicians need to learn how to cope with online limitations. Help might come from the real world of music-making. There is also inherent audio latency in live music. The distance between choir, soloists, and orchestra musicians from stage to orchestra pit in opera imposes a latency of over 50ms, which would make synchronizing impractical without the conductor (Rottondi et al., 2016). Sawchuk's (2003) experiments with asynchronous (a cello masterclass) and synchronous (two duets, with and without an audience) virtual interactions demonstrated that networked playing can be comparable to the sound delay found during live orchestral performances in large concert halls (10-50ms for audio). In sum, humans can adapt to latency. The thresholds for musicians' latency tolerance are (a) optimal, 20-30ms, (b) tolerable, up to 50ms, (c) painful, 50-100ms, and (d) impossible, more than 150ms (Kapur, 2005; Rottondi et all., 2016; Zimmermann et al., 2008). For training purposes, music conducting teachers might relate the delay musicians experience to the ones experienced on stage to network latency in a virtual music teaching context.

Researchers are questioning the physicality, viability, and worthiness of novel systems for NMP (Kapur et al., 2005). Technology innovators are experimenting and presenting systems that aim to approximate network performance to a live music performance. A multi-channel, multi-modal monitoring and recording network system that can capture, store, and distribute music performance events remotely can benefit the teaching of music conducting (Gu et al., 2005). Imagine a conducting class where the instructor, the musicians, and the conducting students, all from different locations, can participate, record, and perform together in a virtual conducting class. There is technological ground yet to be covered until NMP is fully realized, but from a music pedagogy perspective, remote instruction will only benefit from system developments after they become commonplace and economically affordable. Next, I discuss trends, experiences, and perspectives found in research about the use of technology in music conducting.

3. Music Conducting and Technologies as Pedagogical Tools

Music is a gestural-technical and aural communication tool that is learned through practical application. Similarly, conducting is a performance skill with a set of acquired knowledge. Conducting skills include the ability to effectively lead a group of musicians through expressive gestures, facial expressions, body movements, and verbal communication. The knowledge that conductors gather during their preparation reflects their understanding of the score to be conducted according to an interpretation of the composer's intention, the form of the music, and the technique each musician applies to play their specific part. Conducting pedagogy encapsulates best practices to teach such skills to conducting students. Thus, if technology can extend the reach of music-conducting teaching and learning, this would serve music-conducting pedagogues and students.

Technological advances have been made by simulating the display of tempo and dynamics in conducting gestures through digital interfaces. Digital devices might give novice conductors practice opportunities when an actual music ensemble is not available. For example, when articulating with their bows, string performers interpret the arm movements of conductors (Bevilacqua et al., 2007). The use of musical gestures facilitated by technology software might be a tool to teach string students how to articulate sounds by transferring the virtual interface of arm moving to the practice of bowing. Similarly, researchers have suggested situations in which teaching and learning music-conducting with more recent technologies might be explored as pedagogical tools (Bauer, 2016; Bevilacqua et al., 2007; Marrin, 1996; Rottondi et al., 2016). Conducting instructors should investigate new approaches–such as remote teaching, virtual human interaction, or new human-machine interfaces–for teaching conducting to novice conductors.

3.1 Conducting Gestures, Tradition, and Technology Interfaces

Gestures are a quintessential element of human-to-human communication in music conducting. Morrison and Silvey (2022), and Orman et al. (2017) have discussed the use of gestures and how they connect traditional conducting practices. Their research indicates that conducting knowledge is a combination of technical skills, aural image, and kinesthetic awareness–the ability to use the body to communicate musical intent. Music conductors use physical gestures, torso movement, and eye contact as forms of nonverbal communication (Orman et al., 2017). Also, conducting gestures is an interaction between a mental sound image of the score and the sound produced by the players, used at the moment to develop rehearsal strategies, modify instruction, assess, and respond to students and sound stimuli, and formulate concepts of sounds (Morrison & Silvey, 2022). In other words, music conducting is a way of responding to sound stimuli while listening and assessing sound production when it occurs.

Specifically, why are gestures so important in music conducting? Morrison and Silvey (2022) explained that gestures might assist performers' understanding of the music being performed because different aspects of movement are discernible to observers. And what are those aspects? They specify that expressive gestures vary in speed, size, and acceleration and that such gestures are used to inform the form, time signature, articulation, and expressive intent of the music such as phrasing and dynamic variations. These are important concepts to help clarify which aspect of conducting a particular technological tool may assist with when assessing future pedagogical impacts.

Research in capturing performers' gestures through sensors and mapping them to produce computer-synthesized sound is represented by the New Interfaces for Musical Expression (NIME) field (Tanaka & Altavilla, 2012). In the NIME field, researchers have developed systems to gain insight into the notion of MA during human-computer music interaction. Fazekas et al. (2013) developed the Mood Conductor, a system that allows the audience to interact with stage performers to create directed improvisations. Participants act as conductors by communicating emotional intentions to the performers through the web-based smartphone-friendly Mood Conductor app. Emotions are represented by colored blobs in a two-dimensional space (i.e., vertical dimension: arousal or excitation; horizontal dimension: valence or pleasantness). Performers receive the audience's directions via this visual feedback system and improvise performance accordingly in real-time. Such a study represents a research test on the partial substitution of the conductor's role because the audience directly influences the music performance through the manipulation of technological applications.

Music conducting pedagogy regarding the use of the body and gestures to teach musical concepts through technological means has been investigated. Bevilacqua et al. (2007) believe that "physical gesture is a central element for performance but also for the embodiment of music concepts and theory" (p. 124). According to the phenomenon called "kinaesthetic empathy," one can embody others' movements themselves (Koivuen & Wennes, 2011). The researchers explain this further:

> all of our five senses interact so that the contribution of each becomes indistinguishable in the total configuration of perception. Thus, perception concerns the whole sensing body. The unification of the senses comes about through their ongoing integration into a synergic system. This *synaesthetic* system rules our body, but we are unaware of it because we believe in the mechanistic view that we perceive things through the separated channels of perception: seeing by eyes, hearing by ears, and so on. (Koivuen & Wennes, 2011, p. 63)

Therefore, if conducting gestures are perceived by an integrated perception system, perhaps conducting gestures might be used to teach musical concepts through cognitive perceiving empathy in a specific form of "show how" using bodily gestures. The kinaesthetic empathy phenomenon (Koivuen & Wennes, 2011) of embodying others' movements needs to be studied as it relates to the possible connection between conducting gestures and the teaching of musical terms and ideas. Schnell et al. (2011) executed a study around this idea involving gestural control and re-embodiment of recorded sound and music systems. They tested a device in between complex platforms and a gaming controller called *MO,* a modular motion capture device that if hand-held might be used as an augmented object such as a baton. In the experiment, the developer team demonstrated that the MO wireless sensor module may be hand-held, attached to the body, or used to augment objects such as a ball or a kitchen utility to create new instruments and playing techniques. Those new musical playing metaphors enlarge the possibilities for virtual orchestra conducting.

Conducting instructors should consider investigating new approaches for teaching conducting to future music educators. It is important for conducting instructors to be aware of the full capabilities of technological tools that can help conducting students to better perform and understand music conducting. For example, Marrin (1996) described the development of a device called the Digital Baton, which sends data on position, orientation, acceleration, and surface pressure to an external tracking unit. She detailed that an accompanying software system would process the gestural data and allow the Digital Baton user to conduct a computer-generated musical score.

Sarasúa and Guaus (2014a, 2014b) described the role of intuition in understanding the effect of gestural movements on sound production, writing that untrained participants intuitively associated higher raised and more energetic hand movements with loud parts of an orchestra recording. Furthermore, Southerland (2019) argued that different uses of conducting gestures in history also can point to new pedagogical interventions. Hence, researchers have indicated the possibility of a more naturalistic approach to the teaching of conducting that might integrate technological developments in the area.

With regard to technological advances in music education and conducting pedagogy, the use of video recordings has been essential (Grashel, 1991; Silvey & Major, 2014). Grashel (1991) indicated that fundamental conducting competencies could be demonstrated and observed through video. He suggested that baton grip, beat pattern, style, use of left hand, cuing, dynamic indicators, tempo stability, cutoff gestures, and eye contact could be taught through videotaped modules. The premise was that the same content, sequence, and organization of instruction could be distributed equally to all students, not to mention that compared to traditional conducting teaching settings–with the orchestra, videotaped modeling saves time and uses less space.

Once computers progressively gained faster processing power, there was an increased interaction between machines and humans (Ansari et al., 2018). One of the challenges with the use of interactive technologies is the aim for naturalism and intuitive human-computer connections. The constant development of smartphones is the most noticeable and recent example. The use of technology in the realm of music conducting might be organized as such: (a) virtual-robotic orchestras and conductors; (b) gestural controlling systems; and (c) computer-based conducting analytical systems. In the next section, I synthesize experiments in those areas by discussing their contributions for eventual use in conducting pedagogy contexts.

3.1.1. Virtual-Robotic Orchestras and Conductors

A prerequisite for the development of music performance skills is the availability of the instrument. In typical circumstances, music performers have their instruments at their disposal. Even large instruments, like the piano, are available in practice rooms of music schools for students to use. One of the challenges for conducting students is the lack of their "instrument" for practicing–an ensemble. Because of that, part of conductors' preparation is the development of their "inner hearing" so they can embody conducting movements only by looking at a score, without the actual presence of an ensemble to lead. The lack of an instrument for conductors might be mitigated with the advent of virtual orchestras. Advances in virtual reality (VR), sound synthesizing, gesture recognition, and human-computer interfaces have made it possible to design systems that partially fulfill the vacuum of the orchestra as an instrument for pedagogical use (Ferguson & Wanderley, 2010; Schertenleib et al., 2004).

Visual interfaces are the object of testing for the development of virtual orchestras. Sarasúa et al. (2019) designed a system whereby a digital instrument would function as an orchestra that followed the movements of the conductor in real time. Results suggested that the system was usable for providing intuitive—in the sense that the users learned to use the system while performing—control over both loudness and timing. The researchers emphasized this aspect of intuitive learning because it broadens its use for the public, not only for trained music conductors.

Another type of virtual orchestra is the use of a multimodal system with 3D sound and computer interface (Shertenliebe et al., 2004). A magnetically tracked baton is used to control the performance without specific conducting patterns. It is a semi-immersive virtual reality experience–no head-mounted display (HMD) or data gloves–based on a projection screen where virtual animations of musicians play while adapting their performance to the user's gestural tempo, dynamics, and hand and arm movements. The music is pre-codified in MIDI files. The Personal Computer (PC) server controls 3D scenes and sound, processing the gesture input from the magnetic baton, and animating the 3D models accordingly.

In this virtual orchestra system, the sound information in MIDI files is a set of several recordings of the same musical event with variations in dynamics and sections of the music. The sound reflects and blends with the animation of the virtual orchestra. Selected notes within pre-recorded music scores trigger the event, allowing for the manipulation of the animations. Importantly, the system is managed through sound events, not images, avoiding delay in the sound processing because "human senses are more affected by sound distortion then by visual artifacts" (Schertenleib et al., p. 47). The layout of the architecture design of the virtual orchestra breaks down the complexity of an apparent simple behavior of a conductor which is to control for tempo and dynamics of an ensemble performance. Note that these are only two musical parameters amongst an array of musical aspects that a conductor must encompass during ensemble performance and rehearsal.

A third example of virtual orchestra is Ferguson and Wanderley's (2010) Digital Orchestra (DO). This interdisciplinary three-year project relied on different components to build their DO. A PC was used as the synthesis engine. Hardware created for the project functioned as the gestural control surface. Finally, the Mapper–software developed specifically for the project–performed the mapping between gestural data and synthesis parameters such as the information about the baton grip or the amount of energy used by the performer which are captured by sensors and synthesized by the Mapper. The final product was a concert performance of the DO, with music specially composed for the event.

Music in a Universal Sound Environment (MUSE) is a system that captures conductors' musical gestures to drive a MIDI-based music generation system allowing a human user to conduct a fully synthetic orchestra (Carthen et al., 2018). MUSE allows users to control the tempo of a virtual orchestra through basic conducting patterns used by conductors in real-time. Their system also aims to further improve a conductor's technique in an interactive environment. The researchers described how the system facilitates learning through an intuitive graphical interface, and how they utilized techniques from machine learning to process inputs from a low-cost sensor that estimates the beat patterns that a conductor suggests by interpreting hand motions.

In addition to the advent of digital orchestras, conductors themselves have also been replaced virtually. Nijholt et al. (2008) presented the *Virtual Conductor*, a system that performs real-time audio analysis of music played by musicians and uses this analysis to animate a virtual conductor. The system detects the tempo and the dynamics of the music, compares the results with the music score, and changes the conducting behavior when the musicians deviate from the desired performance. The virtual conductor is projected on a large screen in front of the orchestra. Algorithms determine what sounds the musicians are playing and how well. The system has a pedagogical intent, as it was designed to train musicians to follow conductors' gestures. The novelty of the system is that the virtual conductor adapts his movements to give corrective feedback to the musicians. Aside from virtual orchestras and conductors, developers are working on interactive systems aiming to control specific aspects of conducting, usually for the pedagogical purpose of practicing gestural techniques.

3.1.2. Gestural Controlling Systems

In the search for a naturalistic interaction between computers and humans, researchers and technology developers typically select some aspects of music conducting (e.g., tempo, dynamics) to control. There are two ways of human-computer interaction for gestural controlling systems: camera-based systems (i.e., distant cameras and computer vision techniques) and sensor-based systems (i.e., motion sensing devices located in users' hands and body) (Argueta & Chen, 2009; Carôt & Schuller, 2011; Cosentino et al., 2012; Sarasúa et al., 2016; Toh et al., 2013). What differentiates music-conducting electronic systems is the use of one of these two systems. Does the system use contactless devices or close-to-the-body sensing technology? Next, I discuss systems and how they deal with the challenges of emulating the control conductors have over performing ensembles via human-computer interaction.

One such system was designed by Sarasúa et al. (2014b) to learn conductors' gesture variations for gesture-based interaction. The system was used as an interaction technique in a music conducting scenario where gesture variations drove music dynamics. A model was used to allow the user to configure the system by providing variation examples. The system performance and the influence of user musical expertise were evaluated in a user study, which showed that the model was able to learn idiosyncratic variations, thus allowing users to control dynamics, with better performance for users with musical expertise.

Carôt and Schuller (2011) tested a telematic visual-conducting system that focused on the transmission and representation of the conductor's timing instead of their gestures. The idea was to use a computer mouse as a baton to poll the mouse's x and y coordinates which are interweaved with the actual audio stream. Musicians performed remotely according to visualized mouse data representing the tempo of the music. Thus, there is no video of the conductor's gestures and no jitter resulting from the sum of audio plus video streaming in roundtrip transmission. Cosentino et al. (2012) described a human gesture recognition system developed to enable a robot instrument player to recognize the variations in tempo and in articulation dictated by a conductor's movements. The researchers suggested that the interaction ability would allow the partner musicians, as well as the conductor, to appreciate a joint musical performance because of the naturalness of the interaction. In addition, the possibility for the robot to change its performance parameters according to the conductor's directions, thus being synchronized with all the other human musicians, would lead to an improvement in the overall musical performance.

Sarasúa et al. (2016) tested a machine learning (ML) system of personal gesture variations used in music conducting. The idea was to deliver control of music articulation through personal variations of gesture execution using a computer probabilistic model. Participants used a computer mouse and *Microsoft Kinect* as input devices for gesture and articulation. The participant controls a MIDI-performed melody using turned figure-eight gestures with the mouse on legato, tenuto, and staccato styles. Each set of data represents a different articulation of the same gesture which is then paired with outputs encoding the articulation using the probabilistic model. The results were that the model learned the articulations, embedding the user's own expressive gestural qualities. Besides conducting gesture ML, the articulation model offers audio and video representations for user feedback. The authors explained that audiovisual feedback is an action-perception cycle that might assist users in adapting their gestures to target tasks such as the articulation of the melody, tempo control, and dynamics. This sensor-motor learning represents a new pedagogical approach with possible future research applications. Also using *Microsoft Kinect*, Toh et al. (2013) proposed a real-time interactive system that would overcome the limitation of *Kinect* being designed for large body movements, so refined delicate conducting gestures could also be correctly recognized with high levels of accuracy and low latency. The authors' proposed idea is that their interactive conducting system could mediate the interpretation of a conductor's gesture, thus serving as a training tool for students, and for professional conductors and composers as a low-cost experience to shape music virtually.

Orman et al. (2017) tested an augmented immersive Virtual Reality Learning Environment (VRLE) system to increase eye contact, torso movement, and attention skills in novice conductors. Ten music majors tested VRLEs with head tracking or VRLEs without head tracking, including a control group (no VRLEs) in pre- and post-test sessions with a music score. The components of the VRLEs were photos taken of an ensemble from the conductor's perspective to create the virtual environment and instrumental grouping of focused attention outlined with red boxes around the ensemble members. The red box would appear around each focus group that needed the conductors' attention. Full ensemble parts had no red box. Results suggested that students using the augmented immersive VRLE with head tracking had more conducting skill improvement than those not using virtual reality.

Although some researchers have concentrated on gesture-controlling systems, others focused on the computer-based analysis of movement for pedagogical purposes. In typical university conducting classes, the student conducts the ensemble, and the conducting teacher gives feedback. Technology developers are working on systems that might offer new instruments to assist in the analysis and feedback regarding music conducting. Areas of development in conducting pedagogy and technology are movement analysis and gesture-capturing technology (Bevilacqua et al., 2007). In the next section, I report on studies involving computer-based analysis systems.

3.1.3. Computer-Based Conducting Analytical Systems

Movement and gesture analyses have become important aspects in the development of interfaces for gaming platforms and mobile devices (Schnell et al., 2011). In the area of music conducting, researchers have analyzed conductors' nonverbal behaviors. As a compliment to the traditional master-apprentice conducting relationship, researchers have worked in visual, aural, tactile, and multi-sensory feedback modalities to assist conducting instructors with their teaching strategies (Chin-Shyurng et al., 2019; Sarasúa & Guaus, 2014b; Schramm et al., 2015). For example, Sarasúa and Guaus (2014b) proposed a study to analyze (using simple motion capture descriptors– position, velocity, acceleration, jerk, quantity of motion, and contraction index) how participants move when asked to conduct with classical music excerpts, focusing on the influence of the dynamics in the performance. Participants were asked to conduct three classical music fragments while listening to them and were recorded with a commercial depth-sense camera. Results indicated that different tendencies could be found and exploited to design models adjustable to the expectations of the users to control the expressive aspects of a performance by imitating the gestures of conductors.

Chin-Shyurng et al. (2019) proposed a real-time musical conducting gesture recognition system using a camera to capture image inputs and establish a real-time computer-human interaction.

The *Kinect* software development kit created a skeleton model by capturing the palm position of the conductor. Different palm gestures were collected to develop training templates for music conducting. The Results were that the dynamic time-warping algorithm successfully recognized the different conducting gestures at various conducting speeds and achieved real-time dynamic music conducting gesture recognition.

Another example involving music-conducting computer analysis systems is a tool created by Schramm et al. (2015) to aid in the study of basic conducting gestures, also known as meter-mimicking gestures. It is based on the automatic detection of musical metrics and their subdivisions by analysis of hand gestures. Musical metrics are represented by visual conducting patterns performed by hands. These patterns are recognized and evaluated using a probabilistic framework based on Dynamic Time Warping (DTW). DTW is used for measuring similarities between two temporal sequences at different speeds, a method of comparing sequences of time series data, usually based on data from accelerometers worn or used by a person. In conducting, the method is used to compare the recorded gesture of a conductor to a template of standardized gestures in a computer (O'Rourke & Madden, 2011). In Schramm et al. (2015) study, a new metric is proposed for the DTW, allowing better alignment between two distinct gesture movements. The time precision of the conducting gesture is extracted directly from the warping path and its accuracy is evaluated. The authors indicated that the classification scheme represents an improvement over similar approaches.

Armitage et al. (2013) presented *mConduct*, an interactive multimedia system that captures and analyzes conductors' gestures to offer visualization, sonification, and vibrotactile feedback. An accelerometer, a gyroscope, and a magnetometer attached to a baton measure the strength, intensity, and direction of the conductor's gestures. This 3-part measurement system delivers a representation of the baton's movements. The system can stream gestural information which can then be broadcasted and received by computers for processing, analysis, and multi-modal feedback. Visualization software creates a 3D sculpture of the user's gestures. The authors suggested that the virtual avatar encompasses the structure, expression, tempo, and time signature of the music. Because people remember information better when it is represented and learned both audibly and visually (Leavitt, 2012), applications for conducting pedagogy might include building mental models by directing attention to visual information regarding conducting techniques. Collected digital data and visualizations can be stored. The retrieved data might help students and instructors to analyze how the piece of music was shaped.

Chin-Shyurng (2019) used a real-time video gesture recognition system to foster conducting students' virtual practicing. She employed the *Kinect SDK* (Software Development Kit) to capture images of conductors and the DTW (Dynamic Time Warping) algorithm to recognize and classify the conductor's gestures. The study included promising new perspectives for conducting pedagogy with the use of remote technology such as Kinect SDK and DTW for the virtual practice of conducting gestures.

3.1.4. Possible Advantages of Music Conducting Systems

Music Conducting systems–hardware-software applications that electronically emulate some aspect of music conducting gesture expressions for the purpose of technique analysis or remote controlling of robot instruments or virtual orchestra–could be used to rehearse technical, less creative, and non-expressive aspects of a performance or be used to explore facets of ensemble playing in a school setting, with children being allowed to have some level of autonomy. Student conductors could experiment with ways of conducting a passage while playing along with their instrument, putting themselves in both positions simultaneously. Time in front of a virtual ensemble could be almost unlimited (Orman et al., 2007). In combination with a virtual orchestra, these systems could detect mistakes and show better ways of conducting (Nijholt et al., 2008).

Besides simulating orchestral conducting, interaction paradigms–if the system is set to follow the conductor or to record the gestures for analysis–can be used to understand conducting gestures as a continuous process where the gestures might be freely chosen. Moving away from already established conducting patterns, gestural movement analysis can be adjusted to different pedagogical purposes such as instrument performance and the understanding of music theory concepts. Additional pedagogical uses might include breathing and understanding music structure such as phrasing and cadences. Ultimately, interactive conducting systems might assist with technique awareness, interpretation, and expression of music (Bevilacqua et al., 2007). Examples include practicing the effect of gesture on tempo according to arm speed and self-observed movement technique with captured and computer-processed images of conducting gestures, going beyond traditional video recordings. Although computer-based music conducting systems offer new pedagogical possibilities, such a complex human activity as performing, teaching, and learning music conducting presents great challenges until widespread use of these system prototypes is available in teaching settings.

There are few investigations about the impact of new technologies on conducting pedagogy. I found three studies that investigated the effect of technology-based practice on skill achievement in novice conductors (Hollinger & Sullivan, 2007; Sarasúa & Guaus, 2014a; Sarasúa & Guaus, 2014b). Holinger and Sullivan (2007) listed technology tools that demonstrated positive results: (a) a program to aid in score study using a computer-assisted program (Hudson, 1996); (b) the Music Conducting Trainer (MCT) to aid in the steadiness and shape of participants' conducting patterns in four different modes while music followed in the tempo of the beat pattern of participants (Schwaeger, 1984); and (c) a comparison between Marrin's Digital Conducting System (DCS) and Max Matheu's Radio Baton (RB) (Kraus et al., 2004). In general, results showed that computer feedback was effective in improving selected fundamental skills such as preparatory ictus, release gestures, and the learning of legato skills. Unfortunately, Hollinger's analysis is already outdated, as are experiments completed in the 1990s.

Somewhat recently, Sarasúa and Guaus (2014b) conducted a computational comparative study on dynamics in music conducting. Participants conducted pre-recorded fragments of Beethoven's Eroica using a musical interface conductor metaphor–imitation of conductor's gestures–of audio, multi-perspective video and aligned music score. A virtual skeleton that included 15 joints captured and recorded the motion of 24 participants aligned with the audio-video and Eroica score. Every joint had position, velocity, acceleration, and jerk–a derivative of acceleration–recorded. The movements served as predictors and the loudness of the music served as the independent variable. Linear regression models indicated that motion capture descriptors correlated to loudness differently among participants. For instance, in loud parts, most participants had greater amplitude movements, but some just raised their hands higher. The authors suggested that although it was very difficult to obtain a model that was generalized to all participants, different tendencies might help developers to design tailored models to teach conducting using computational analysis.

More recently, Kshirsagar and Annadate (2017) compared ten music-conducting gestural recognition systems using Dynamic Time Warping (DTW). The authors found both pedagogical advantages and technological challenges in each of the ten gestural recognition systems. Gestural interaction systems might substitute partially for conducting real music ensembles and enhance the teaching experience. However, I could find no documented evidence on the widespread use or large-scale tests on these systems, especially in educational settings.

CONCLUSION

The use of technology in music conducting contexts might help music teachers and pedagogues to consider more thoughtfully how students can be taught conducting using new approaches. When new technologies come along in some aspect of human life––for example when in the job market a machine substitutes for the labor of a human being––human intelligence and inventiveness cause us to reflect on what is essentially human in social life. That philosophical problem should be considered when conducting instructors reflect upon music conducting coursework and the use of technology as a potential pedagogical tool. We have not found a substitute for human-to-human interaction when it comes to education or the teaching of music conducting. Nevertheless, investigating how technology can assist music conducting teachers in the teaching of conducting virtually, or when a music ensemble is not available for a conducting student, might inform and possibly enhance the pedagogy of conducting. The use of technology to teach and learn conducting needs to be further studied to guarantee that such a traditional artistic and educational practice remains relevant in the ever-developing technological world.

REFERENCES

AKTÜRK, A. Ata; DEMIRCAN, Ozlen. A review of studies on STEM and STEAM education in early childhood. **Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi (KEFAD)**, v. 18, n. 2, p. 757-776, 2017. Available at: < https://www.researchgate.net/publication/319702309_A_Review_of_Studies_on_STEM_and_STEAM_Education_in_Early_Childhood?enrichId=rgreq-1d8e2ead343df01a3b5e5590998057a6-XXX&enrichSource=Y292ZXJQYWdlOzMxOTcwMjMwOTtBUzo1MzkyMTMxNzY4NzI5NjBAMTUwNTU2OTg0NzU5Ng%3>. Access on: 2 aug. 2021.

ANSARI, Fazel; EROL, Selim; SIHN, Wilfried. Rethinking human-machine learning in industry 4.0: how does the paradigm shift treat the role of human learning? **Procedia manufacturing**, v. 23, p. 117-122, 2018. Available at: <https://doi.org/10.1016/j.promfg.2018.04.003>. Access on: 1 may. 2022.

ARGUETA, Carlos Rene; KO, Ching-Ju; CHEN, Yi-Shin. Interacting with a music conducting system. In: **Human-Computer Interaction. Novel Interaction Methods and Techniques: 13th International Conference, HCI International 2009, San Diego, CA, USA, July 19-24, 2009**. Springer Berlin Heidelberg, Proceedings, Part II 13, p. 654-663, 2009. Available at: < https://rdcu.be/dTqkA>. Access on: 8 jan. 2022.

ARMITAGE, Joanne; NG, Kia. mConduct: a multi-sensor interface for the capture and analysis of conducting gesture. In: **Electronic Visualisation in Arts and Culture**. London: Springer London, 2013. p. 153-165. Available at: <https://link.springer.com/chapter/10.1007/978-1-4471-5406-8_11>. Access on: 8 jan. 2022.

BAUER, William I. Technological affordances for the music education researcher. **Update: Applications of Research in Music Education**, v. 34, n. 3, p. 5-14, 2016. Available at: < https://doi.org/10.1177/8755123314568570>. Access on: 10 jan. 2022.

BENAQUISTO, Lucia; GIVEN, Lisa. The SAGE encyclopedia of qualitative research methods. **Given L, ed**, v. 413, 2008. Available at: <http://dx.doi.org/10.4135/9781412963909.n48>. Access on: 13 sept. 2021.

BOWEN, Glenn A. Document analysis as a qualitative research method. **Qualitative research journal**, v. 9, n. 2, p. 27-40, 2009. Available at: <https://doi.org/10.3316/QRJ0902027>. Access on: 6 aug. 2021.

BEVILACQUA, Frederic et al. 2007: Wireless Sensor Interface and Gesture-Follower for Music Pedagogy. **A NIME Reader: Fifteen Years of New Interfaces for Musical Expression**, p. 267-284, 2017. Diponível em: < https://doi.org/10.1007/978-3-319-47214-0_18>. Access on: 15 feb. 2022.

BRAUN, Virginia; CLARKE, Victoria. Using thematic analysis in psychology. **Qualitative research in psychology**, v. 3, n. 2, p. 77-101, 2006. Available at: <https://doi.org/10.1191/1478088706qp063oa>. Access on: 16 aug. 2021.

BRADER, Andy. Synchronous learner support for music-sequencing software. **Journal of Music, Technology and Education**, v. 2, n. 2-3, p. 159-174, 2009. Access on: 05 feb. 2021.

CARÔT, Alexander; SCHULLER, Gerald. Towards a telematic visual-conducting system. In: **Audio Engineering Society Conference: 44th International Conference: Audio**

Networking. Audio Engineering Society, 2011. Available at: < https://aes2.org/publications/elibrary-page/?id=16137>. Access on : 5 feb. 2022.

CARTHEN, Chase D. et al. MUSE: A Music Conducting Recognition System. In: **Information Technology-New Generations: 14th International Conference on Information Technology**. Springer International Publishing, 2018. p. 363-369. Available at: <https://doi.org/10.1007/978-3-319-54978-1_49>. Access on: 5 feb. 2022.

CHIN-SHYURNG, Fahn; LEE, Shih-En; WU, Meng-Luen. Real-time musical conducting gesture recognition based on a dynamic time warping classifier using a single-depth camera. **Applied Sciences**, v. 9, n. 3, p. 528, 2019. Disponível em: <https://doi.org/10.3390/app9030528>. Access on: 15 mar. 2022.

CORBIN, Juliete., & SRAUSS, Anselm. (2008). **Qualitative research**. *Techniques and procedures for developing grounded theory*, *3*. Available at: < https://toc.library.ethz.ch/objects/pdf/z01_978-1-4129-0644-9_01.pdf>. Access on: 22 aug. 2021.

COSENTINO, Sarah et al. Music conductor gesture recognition by using inertial measurement system for human-robot musical interaction. In: **2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)**. IEEE, 2012. p. 30-35. Available at: <https://doi.org/10.1109/ROBIO.2012.6490939>. Access on: 28 feb. 2022.

DILLMANN, Rüdiger. Teaching and learning of robot tasks via observation of human performance. **Robotics and Autonomous Systems**, v. 47, n. 2-3, p. 109-116, 2004. Available at: <https://doi.org/10.1016/j.robot.2004.03.005>. Access on: 06 mar. 2022.

FAZEKAS, György; BARTHET, Mathieu; SANDLER, Mark B. Mood conductor: emotion-driven interactive music performance. In: **2013 Humaine Association Conference on Affective Computing and Intelligent Interaction**. IEEE, 2013. p. 726-726. Available at: <https://doi.org/10.1109/ACII.2013.165>. Access on: 03 mar. 2021.

FERGUSON, Sean; WANDERLEY, Marcelo M. The McGill Digital Orchestra: An Interdisciplinary Project on Digital Musical Instruments. **Journal of Interdisciplinary Music Studies**, v. 4, n. 2, 2010. Available at: <https://www-archive.idmil.org/_media/publications/2010/ferguson_2010_jims.pdf>. Access on: 05 feb. 2021.

GAVER, William W. Technology affordances. In: **Proceedings of the SIGCHI conference on Human factors in computing systems**. 1991. p. 79-84. Available at: <https://dl.acm.org/doi/pdf/10.1145/108844.108856>. Access on: 06 mar. 2022.

GILLIES, Marco et al. Human-centred machine learning. In: **Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems**. 2016. p. 3558-3565. Available at: <https://doi.org/10.1145/2851581.2856492>. Access on: 06 mar. 2022.

GODØY, Rolf Inge; JENSENIUS, Alexander Refsum. **Body movement in music information retrieval**. 2009. Available at: < https://www.researchgate.net/profile/Alexander-Jensenius/publication/220723329_Body_Movement_in_Music_Information_Retrieval/links/57ed2e4b08ae7be720544b13/Body-Movement-in-Music-Information-Retrieval.pdf> Access on: 12 apr. 2022.

GLASER, Barney et al. Strauss.(1967). **The discovery of grounded theory**. Alsine de Gruyter, New York, 1967. Access on: 15 sept. 2021.

GRASHEL, John. Teaching Basic Conducting Skills Through Video. **Music Educators Journal**, v. 77, n. 6, p. 36-37, 1991. Available at: <https://doi.org/10.2307/3398211>. Access on: 22 mar. 2022.

GU, Xiaoyuan et al. Network-centric music performance: practice and experiments. **IEEE Communications Magazine**, v. 43, n. 6, p. 86-93, 2005. Available at: <https://doi.org/10.1109/MCOM.2005.1452835>. Access on: 17 mar. 2022.

HOLLINGER, Diana; SULLIVAN, Jill M. The Effects of Technology-Based Conducting Practice on Skill Achievement in Novice Conductors. **Research and Issues in Music Education**, v. 5, n. 1, p. 1-6, 2007. Available at: <https://eric.ed.gov/?id=EJ814930>. Access on: 02 apr. 2022.

HOLUB, Petr et al. UltraGrid: Low-latency high-quality video transmissions on commodity hardware. In: **Proceedings of the 20th ACM international conference on Multimedia**. 2012. p. 1457-1460. Available at: < http://doi.acm.org/10.1145/2393347.2396519 >. Access on: 10 apr. 2022.

HUDSON, Mark Edward. **The development and evaluation of a computer-assisted music instruction program as an aid to score study for the undergraduate wind band conducting student**. University of Florida, 1996.

KAPUR, Ajay et al. Interactive Network Performance: a dream worth dreaming? **Organised Sound**, v. 10, n. 3, p. 209-219, 2005. Available at: <http://dx.doi.org/10.1017/S1355771805000956>. Access on: 04 mar. 2022.

KOIVUNEN, Niina; WENNES, Grete. Show us the sound! Aesthetic leadership of symphony orchestra conductors. **Leadership**, v. 7, n. 1, p. 51-71, 2011. Available at: <https://doi.org/10.1177%2F1742715010386865>. Access on; 28 feb. 2022

KRAUS, B. N., GONZALEZ, G. M., HILL, G. W., & HUMPHREYS, J. T. Interactive computer feedback on the development of fundamental conducting skills. **Journal of Band Research**, v. 39, n. 2, p. 35-44, 2004.

KUCKARTZ, Udo. **Qualitative text analysis: A guide to methods, practice and using software**. Sage, 2014. Available at: <https://books.google.com.br/books?hl=pt-BR&lr=&id=9B2VAgAAQBAJ&oi=fnd&pg=PP1&dq=Kuckartz,+U.+(2014).+Qualitative+text+analysis:+A+guide+to+methods,+practice+and+using+software.+Sage.&ots=w2CUa8OjhX&sig=WHQ6vCYgEfPCqaNFu7Gx73r_uXs#v=onepage&q=Kuckartz%2C%2>. Access on: 25 aug. 2021.

KSHIRSAGAR, Ms Snehal; ANNADATE, M. N. Survey on Music Conducting Gestures using Dynamic Time Warping. **Int. Res. J. Eng. Technol**, v. 4, p. 2835-2839, 2017. Available at: < https://d1wqtxts1xzle7.cloudfront.net/53592136/IRJET-V4I5713-libre.pdf?1497954541=&response-content-disposition=inline%3B+filename%3DSurvey_on_Music_Conducting_Gestures_usin.pdf&Expires=1727710005&Signature=YHA6KOoKc1UsAkz8m7eB7yW-QNFyCXyRS43d7J71BcwKcSO~>. Access on: 15 apr. 2022.

LEAVITT, Matthew. Learning from Visuals How Well-Designed and Well-Used Visuals Can Help Students Learn: An Academic Review. **Arizona State University, Web, 5th December**, 2012. Available at: < https://www.scribd.com/document/237395902/Leavitt-a-LEARNING-FROM-VISUALS-How-Well-Designed-and-Well-Used-Visuals-Can-Help-Students-Learn>. Access on: 05 mar. 2022.

MARRIN, Teresa Anne. **Toward an understanding of musical gesture: Mapping expressive intention with the digital baton**. 1996. Tese de Doutorado. Massachusetts Institute of Technology. Available at: < https://cir.nii.ac.jp/crid/1571980075454125824>. Access on: 10 jan. 2022.

MERRIAM, Sharan B. **Qualitative Research and Case Study Applications in Education. Revised and Expanded from" Case Study Research in Education"**. Jossey-Bass Publishers, 350 Sansome St, San Francisco, CA 94104, 1998. Access on: 16 aug. 2021.

MORRISON, S. J.; SILVEY, B. **Conducting**. In G. McPherson (Ed.), Oxford Handbook of Music Performance: Insights from Education, Musicology, Psychology, Science and Medicine, 2022. Available at: <https://doi.org/10.1093/oxfordhb/9780190056285.013.13>. Access on: 15 feb. 2022.

NIJHOLT, Anton et al. The virtual conductor: learning and teaching about music, performing, and conducting. In: **2008 Eighth IEEE International Conference on Advanced Learning Technologies**. IEEE, 2008. p. 897-899. Available at: < https://doi.org/10.1109/ICALT.2008.43>. Access on: 26 mar. 2022.

ORMAN, Evelyn K.; PRICE, Harry E.; RUSSELL, Christine R. Feasibility of using an augmented immersive virtual reality learning environment to enhance music conducting skills. **Journal of Music Teacher Education**, v. 27, n. 1, p. 24-35, 2017. Available at:

<https://doi.org/10.1177/1057083717697962>. Access on: 26 mar. 2022.

O'ROURKE, Conor; MADDEN, Michael. Activity recognition based on accelerometer data using dynamic time warping with ensembles. In: **Proceedings of AICS-2011, the 22 nd Conference on Artificial Intelligence and Cognitive Science**. 2011.

PATTON, Michael Quinn. Two decades of developments in qualitative inquiry: A personal, experiential perspective. **Qualitative social work**, v. 1, n. 3, p. 261-283, 2002. Access on: 10 aug. 2021.

PROSKE, Uwe; GANDEVIA, Simon C. The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. **Physiological reviews**, 2012. Available at: <https://doi.org/10.1152/physrev.00048.2011>. Access on: 05 mar. 2021.

RODET, Xavier et al. Study of haptic and visual interaction for sound and music control in the phase project. In: **Proceedings of the 2005 conference on New interfaces for musical expression**. 2005. p. 109-114. Available at: <https://www.nime.org/proceedings/2005/nime2005_109.pdf>. Access on: 10 may 2022.

ROTTONDI, Cristina et al. An overview on networked music performance technologies. **IEEE Access**, v. 4, p. 8823-8843, 2016. Available at: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7769205>. Access on: 23 apr. 2022.

SARASÚA, Álvaro; URBANO, Julián; GÓMEZ, Emilia. Mapping by observation: Building a user-tailored conducting system from spontaneous movements. **Frontiers in Digital Humanities**, v. 6, p. 3, 2019. Available at: <https://doi.org/10.3389/fdigh.2019.00003>. Access on: 15 jan. 2022.

SARASUA, Alvaro; CARAMIAUX, Baptiste; TANAKA, Atau. Machine learning of personal gesture variation in music conducting. In: **Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems**. 2016. p. 3428-3432. Available at: <https://doi.org/10.1145/2858036.2858328>. Access on: 15 jan. 2022.

SARASÚA, Álvaro; GUAUS, Enric. Beat tracking from conducting gestural data: a multi-subject study. In: **Proceedings of the 2014 International Workshop on Movement and Computing**. 2014. p. 118-123. Available at: <https://doi.org/10.1145/2617995.2618016>. Access on: 15 feb. 2022.

SARASÚA, Álvaro; GUAUS, Enric. Dynamics in Music Conducting: A Computational Comparative Study Among Subjects. In: **NIME**. 2014. p. 195-200. Available at: <http://www.phenicx.upf.edu/system/files/publications/464_paper.pdf>. Access on: 25 jan. 2022.

SAWCHUK, Alexander A. et al. From remote media immersion to distributed immersive performance. In: **Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence**. 2003. p. 110-120. Available at: <http://doi.acm.org/10.1145/982484.982506>. Access on: 18 feb. 2022.

SCHERTENLEIB, Sebastien et al. Conducting a virtual orchestra. **IEEE MultiMedia**, v. 11, n. 3, p. 40-49, 2004. Available at: < https://doi.org/10.1109/MMUL.2004.5>. Access on: 30 jan. 2022.

SCHRAMM, Rodrigo; JUNG, Cláudio Rosito; MIRANDA, Eduardo Reck. Dynamic time warping for music conducting gestures evaluation. **IEEE Transactions on Multimedia**, v. 17, n. 2, p. 243-255, 2014. Available at: <https://doi.org/10.1109/TMM.2014.2377553>. Access on: 10 jan. 2022.

SCHNELL, Norbert et al. Playing the "MO"-Gestural Control and Re-Embodiment of Recorded Sound and Music. In: **NIME**. 2011. p. 535-536. Available at: <https://www.nime.org/proceedings/2011/nime2011_535.pdf>. Access on: 30 jan. 2022.

SILVEY, Brian A.; MAJOR, Marci L. Undergraduate music education majors' perceptions of their development as conductors: Insights from a basic conducting course. **Research Studies in Music Education**, v. 36, n. 1, p. 75-89, 2014. Available at: <https://doi.org/10.1177/1321103X14523532>. Access on: 25 jan. 2022.

SNYDER, Hannah. Literature review as a research methodology: An overview and guidelines. **Journal of business research**, v. 104, p. 333-339, 2019. Available at: < https://doi.org/10.1016/j.jbusres.2019.07.039>. Access on: 22 sept. 2021.

SOUTHERLAND, William. Giving music a hand: conducting history in practice and pedagogy. **The Choral Journal**, v. 59, n. 8, p. 30-43, 2019. Available at: < https://www.jstor.org/stable/26601977>. Access on: 10 feb. 2021.

SCHWAEGLER, David Gary. **A computer-based trainer for music conducting: the effects of four feedback modes (CAI, psychomotor)**. The University of Iowa, 1984.
TANAKA, Atau; ALTAVILLA, Alessandro; SPOWAGE, Neal. **Gestural musical affordances**. Sound and Music Computing, 2012. Available at: <http://hdl.handle.net/2086/13580>. Access on: 06 may 2022.

TOH, Leng-Wee; CHAO, Wilber; CHEN, Yi-Shin. An interactive conducting system using kinect. In: **2013 IEEE International Conference on Multimedia and Expo (ICME)**. IEEE, 2013. p. 1-6. Available at: <https://doi.org/10.1109/ICME.2013.6607481>. Access on: 02 mar. 2022.

ZIMMERMANN, Roger et al. Distributed musical performances: Architecture and stream management. **ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)**, v. 4, n. 2, p. 1-23, 2008. Available at: < http://doi.acm.org/10.1145/1352012.1352018>. Access on: 22 mar. 2022.

BIOGRAPHY

Adamilson Abreu has a PhD from the University of Missouri – USA. Artistic achievements include a 3rd place in the Concurso Nacional de Canto "Cidade de Araçatuba" and the finals in the Concurso Nacional de Coros FUNARTE. Mr. Abreu has appeared as choir master for opera productions and conducted orchestras such as the University of Missouri's Philharmonic Orchestra and the Orquestra Sinfônica do Theatro da Paz. Research interest is music conducting, especially on the use of choral singing as pedagogical tool for music education. Mr. Abreu is a full-time tenured faculty at the Universidade Federal do Pará - Brazil. ORCID: https://orcid.org/0000-0003-2365-3170 | Email: abreu@ufpa.br