

# Can gesture recognition in a live digital music performance media essay

[Media](#)



**ASSIGN  
BUSTER**

School of Art & Media Digital Art & Technology Can Gesture Recognition in a Live Digital Music Performance engage the audience as much it engages its user? Christopher Tough 10179279

## **Contents**

### **Abstract**

Chapter 1 Introduction - an overview of Gesture Recognition. Chapter 2 History of Human Computer Interaction (HCI). Chapter 3 Gestures in Live Music. Chapter 4 Live Digital Music & its audience. Chapter 5 Bridging the Gap. Chapter 6 Alternative technologies. Chapter 7 Conclusion

## **Bibliography**

### **Abstract**

Interaction with an audience in a live digital performance is sometimes restricted to a minimum value, this being because of a performer hiding behind their medium, predominantly a computer screen. The main goal of a DJ/live artist is to share and enjoy music through live mixing/performance. Unfortunately, this only allows the audience a partial insight into their interaction with their interfaces. Although, this is cannot be helped in certain aspects of live digital performance, in more recent years, more technologies have been developed to re-engage the audience, allowing the artist to step away from their screens. This dissertation focuses on how Gesture Recognition has been developed for digital performances to help rediscover that engagement the audience once had from their live acoustic counterparts. These live analogue performances used expressive gestures to emotionally engage their audience. With increasing developments in Gesture

Recognition software, these artists can perform gestural movements in order to change parameters on their equipment in a more intuitive way, giving more of an insight to the audience. However do these gestural interfaces being developed for show engage the audience as much as it engages its user?

## Chapter 1

### Introduction – an overview of gesture recognition

Gesture recognition is the interpretation of human motion by a computational device. (Mitra, S. 2007) This means that a computer's computational software containing mathematical algorithms can calculate our bodily motion or state through a series of inputs. Research of gesture recognition allows us to create systems which can identify these specific human gestures and use them to convey information or for device control. G. Kurtenbach says " A gesture is a motion of the body that contains information" (Kurtenbach G. & Hulthen E. A, 1990). Therefore, a gesture can be commonly recognized as a movement which embodies a special meaning. In personal computing, gestures are most often used for input commands. (Turk, M. 2000) Performers have been using sensors to trigger, control and manipulate sounds via computer software for over 50 years, such as the pioneering work by Max Mathews, for example, who use an augmented radio baton to trigger and continuously control a MIDI synthesizer. However it was just after the end of the first world war when the first Gesture Recognition Device was first invented. The 'Theremin'. The theremin is an electronic musical instrument that does not require touch but accepts a user's gesture to create music. (Billingshurst, M. 2011) It was invented by a young Russian <https://assignbuster.com/can-gesture-recognition-in-a-live-digital-music-performance-media-essay/>

physicist Lev Sergeevich Termen or as most people in the western world know him Leon Theremin. The Theremin works using two metal antennas which sense the position of the users hands, which in turn, with one hand you control the oscillators for frequency and the volume with the other. (Glinsky, A. 2000). Recognizing gestures as an input allows computers to be more accessible for the physically-impaired and also makes interaction more natural in several aspects of the digital world. For example, a lot of emerging technologies can cause socially challenging problems; this is because the user has to learn a new GUI (Graphical User Interface) which often contains several new unnatural experiences. However, human communication is a combination of speech and gestures. Gestures are used for everything from pointing at a person to get their attention to conveying information about space." Gestures are part of the non-verbal conversation and are used consciously as well as subconsciously. Gestures are a basic concept of communication and were used by humans even before speech developed", (Nesselrath, R. Alexandersson, J. 2009). What Nesselrath tells us here, is that gestures have the potential to be a huge enrichment to an intuitive human-computer interaction. Even so, with the social challenges mentioned above, gestures must be simple and universally acceptable. Gestures associated with speech are referred to as gesticulation and gestures which function independently of speech are referred to as autonomous. This study of gestures and other non-verbal types of communication are known as "Kinesics". (Birdwhistell. R. L. 1970). The term is said have been coined in 1952 by Ray Birdwhistell, a ballet dancer turned anthropologist who wished to study how people communicate through posture, gesture, stance, and

movement. Kinesics is a topic which has been thoroughly researched and it has been heavily linked with how body language within a musical performance has sufficient dynamic information which is recognized by an audience according to emotional intention performed by an artist. (Salgado, A. 2007) This brings us onto the evolution of computer music and its connection with gesture recognition. There are several reasons to why artists are using gesture recognition in their live performances. For example, there is no single permitted set of options, such as a menu, but rather a series of continuous controls and there is an instant response to the user's movements. These reasons open up a plethora of possibilities and release the restrictions in the artist's performance. Because of the evolution of computers, thanks to the production of inexpensive ubiquitous sensor devices, and exceptionally powerful computer hardware, it has led to a surplus in methods of sound synthesis and allows a mass audience direct access to real-time computer-generated sound. " Computer Music has come to mean two things: the direct synthesis of sound by digital means and computer-assisted composition and analysis." (Baggi, 1991). What Baggi means here is that firstly the music can be played by any digital medium, as of today a major example of a modern composer of computer music are known as Live Artists, whom use programs such as Ableton to compose and play live and pre-recorded sounds. Secondly computer-assisted composition and analysis, is grouped together and not separated because this directly refers to software that can be utilised to aid with composition when the composer is writing and also when a fellow composer wants to study or learn another piece at the same time. Marcelo Wanderley tells us of this growing

maturity in new innovative technologies that have been developed for live digital performance incorporating gesture recognition. " Both signal and physical models have already been considered as sufficiently mature to be used in concert situations". This maturity has led us to discover multiple ways of capturing this data. R Nesselrath explains that there are two ways we are utilising modern technologies for recording gestures. Firstly, " Non-instrumental projects recognize hand and finger postures with cameras and image processing algorithms." (Nesselrath, R. Alexandersson, J. 2009). These hand and also body gestures can be amplified by controllers which can recognize and interpret specific gestures, for example a wave of the hand, for instance, might terminate the program. A prime example is the use of the Microsofts' " Kinect". Kinect is a motion sensing input device which was released in November 2010 and was according to the Guinness Book of Records, the fastest-selling consumer electronics device. (Alexander, L. 2011). There are three innovative working parts to a Kinect sensor, they are, a Colour VGA video camera, a Depth sensor, and a multi-array microphone. This hardware and also the software behind a Kinect is able to detect and track " 48 points on each of the user's body, mapping them to a digital reproduction of that user's body shape and skeletal structure, including facial details". (Crawford, S.) The second way of utilising modern technologies to capture our gestural information are projects which use instruments for recording, for example sensor apparel or hand held devices with integrated accelerometers, gyroscopes or contact microphones. The most obvious example of a hand held device used as an instrument is the Wii Controller, however a more modern technique is that of a MOGEES

(Mosaicing Gestural Surface). According to Bruno Zamborlin, creator of Mogees, " Mogees is a system that allows you to transform any object into a musical instrument just by placing a contact microphone on it." (Zamborlin, B. 2012). The Mogee simply works by connecting a contact microphone to a computer and then using the computer to analyze the audio input signals and extracts the information about how the user is touching the surface. With these mature technologies stated above, a vast majority of artists who perform computer-generated music in popular culture still fail to step away from behind their laptop screens or decks. The use of digital media and technology for performing disguises the physicality of the experience for an audience unlike that which is produced from an acoustic performance which makes music inherently physical. So what is the history between gestures and Human-Computer interaction? What is our understanding of gestures in music? How can the use of gesture recognition help re-establish the disembodiment of music in current digital performances? What new concepts have emerged to bridge this gap? At this moment in time, there are many new styles of digital performance that have broken through thanks to Computer-Generated Music and gesture recognition. This dramatic change in the field of musical composition has been so fast that it has allowed the composer greater flexibility and a richer source of sounds. Current performances include the use of gestures to control devices such as production line robotic arms to play acoustic instruments such as.....(Lost Research on Artist)!

## **Gestures used in Live Music**

A musical gesture can be taken in quite a broad sense. It does not mean only movement, but a movement which can express something. In musical performance, gestures are widely used in different aspects. It is a thoroughly researched field and there are a variety of in depth analysis's however there is a basic distinction for these gestures which are known as physical gestures and mental gestures (Iazzetta. F, 1997). Physical gestures do not only encompass how to produce a sound from an instrument but also how the body moves and its posture while accompanying the sound. For example, the hands are a means of fine action due to their dexterity and the amount of responsive nervous receptors in the finger tips; the feet are mostly suited to the performance of slower and more static movements, whilst other body parts' roles serve as the general support (stability) for the instrument (Cadoz. C, 1988), however there are other roles depending on the instrument. This brings us on to more modern musical gestures. People argue that actions such as turning knobs or pushing levers, which are current in today's technology used by computer musicians or computer music performers or more current live electronic artists, are movements which cannot be considered as gestures. Typing words on computer's keyboard has nothing to do with gesture since the movement of pressing each key does not convey any special meaning." Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed." (Kurtenbach and Hulteen, 1990) However, the situation is completely different when a musician plays something on a piano keyboard: the result,



that is, the musical performance. This is understandable yet electronic artists argue that there are major differences between that of turning a knob at a given moment in time and the pressing a key of computer's keyboard. This is because the modern uses have taken the musical gesture away or the so called instrument away so the performer's body has come to the close attention giving it its own musical gesture. An Interview from renowned site Resident Advisor where an interview with Daniel Brandt from the Brandt Brauer Frick Ensemble discusses the concept of 'Beyond the Laptop', Brandt exerts that there are good live acts that mainly just use a laptop, but the use of a laptop " doesn't really feel live and you can't see, as the audience, what this person's actually doing. The music can sound complicated but then he maybe just switches between different Ableton lines in the arrangement". This establishes the idea that artists which use laptop technologies in musical performance do not produce any musical gesture with embodied information. (Keeling, R. 2012) Conducting is another example of physical gesture as it can be viewed as a way of controlling high-level aspects of performance of multiple instruments with a physical gesture but without direct contact with the instruments. According to Paul Kolesnik's study of Conducting Gesture Recognition, Analysis and Performance System, this example is reinforced through the explanation that there are two principal functions of an orchestral conductor. The first being to " indicate the timing information for the beats of the score in order to synchronize the performance of the musical instruments, and to provide the gestures to indicate his or her artistic interpretation of the performance". (Kolesnik, P. 2004) The second is the introduction of " a degree of variation and personal

interpretation in the musical performance, and is represented by a number of gestures with a high degree of expressivity." (Kolesnik, P. 2004) Mental gestures differ from physical gestures because they are closely related to "the processes of composition, interpretation and listening" (Iazzetta, F, 1997). A composer may use the term musical gestures to "designate a sequence of events within a space of musical parameters". (Wanderley, M, 2000). So mental gestures refer to physical gestures and their relationship and how they occur as an idea or an image of another gesture. An explanation for this is that the mental gesture is learned through the experience and stored into the memory to be used. (Iazzetta, F, 1997). This means that the listener performs the mental gesture while the performer performs the physical gesture. Mental gesture affects the audience/listener through a multitude of perceptual and cognitive mechanisms which have yet to be fully described. Though the human mirror neuron system is a key explanation to how these mental gestures work. The mirror neuron system is said to be "a mechanism allowing an individual to understand the meaning and intention of a communicative signal by evoking a representation of that signal in the perceiver's own brain". (Molnar-Szakacs et al, 2006). What we can understand from is that listeners are able to perceive speech and gestures by a way of articulatory gestures in a way they would perform themselves in order to produce a similar signal. This theory proposes that the brain is able to extract gestural information from the signal. An example of one of these studies is Music-related Motor Learning (Buccino et al., 2004; Calvo-Merino et al., 2004). Buccino's study of the role of the mirror system in motor learning was through the experimentation of the motor system in both human and

non-human primates. In primates, he expresses that a set of neurons are discharged during the " execution of both hand and mouth object-directed actions also respond when a monkey observes another monkey or an experimenter performing the same or similar action." (Buccino et al. 2006). In humans, Buccino strengthens this theory with divulging various methods used to support this notion, for example, experiments with neurophysiological, behavioural, and brain imaging techniques were undertaken. One example Buccino tells of is one at the beginning of the century in 2002, by Craighero and associates. They conducted research in assessing the reaction time of the volunteers. Volunteers were asked to prepare to grasp as fast as possible a bar orientated in either clockwise or anti-clockwise after presenting a picture showing the right hand. Two experiments were carried out. In the first, the picture shown was of a mirror image of the final position in which the hand required to grasp the bar. In the second, two images were produced to the volunteers representing 90 degree rotations of the hand in leftward and rightward directions. Both experiments concluded with subjects having faster responses when delivered with stimuli, however it was the level of similarity between the observed and executed movement led them to a further advantage in the task. (Buccino et al. 2006). These experiments reinforce the fact that mirror neurons exist but they also support that the fact observed actions such as gestures in music can be reflected in the motor representation for the same action of the observers. When it boils down to our mirror-neurons and music Molnar-Szakacs and Overy explain it beautifully." The ability to create and enjoy music is a universal human trait and plays an important role in the daily life of most

cultures. Music has a unique ability to trigger memories, awaken emotions and to intensify our social experiences. We do not need to be trained in music performance or appreciation to be able to reap its benefits already as infants, we relate to it spontaneously and effortlessly." (Molnar-Szakacs et al, 2006)

## **Chapter 2**

### **The growth of HCI**

Where did the term HCI come from? HCI or Human-Computer Interaction is an area of research and practice that came about in the early 1980's. The term " human-computer interaction" is commonly used interchangeably with term such as " man-machine interaction" (MMI), " computer and human interaction" (CHI) and " human-machine interaction" (HMI) but it is predominantly known as HCI. An over simplified definition of HCI might say that it is " the study of the interaction between humans and computers." (Carroll, J. 2013). This can be seen in a general point of view as an acceptable definition but it does not by far justify the actual complexity and multi-disciplinary nature of the subject. In previous decades the majority of computer users were themselves programmers and designers of computer systems. Consequently, a person using a computer system was likely to have been immersed in the same conventions and culture as the individual who designed it. It was not till the mid-1990s did the study of Human-Computer Interaction (HCI) finally take centre stage with the release of Windows 95 bursting upon the scene. Brad Myers, believes that research in HCI has been incredibly successful because of Windows 95 and its " ubiquitous graphical interface", he also mentions that it was based on the Macintosh, which in <https://assignbuster.com/can-gesture-recognition-in-a-live-digital-music-performance-media-essay/>

turn based itself upon the work at Xerox PARC, which in turn again is based upon early research at the Stanford Research Laboratory (now SRI) and at the Massachusetts Institute of Technology (MIT). (Myers, B. 1996). However it was the challenge through the 1970's that started this research and practice. With the emergence of personal computing in the later 1970s, there was eminent change taking place. Personal computing included both personal software (text editors, spreadsheets, and interactive computer games) and personal computer platforms, for example operating systems, and programming languages. This saw a substantial growth in the number users who are not computer users. This change has focused attention upon the needs of what Eason (Eason, K. D. 1976) has termed the naive user and the lack of understanding of the naive user on the part of many designers. It also created awareness to the problems with computers with respect to usability for those users who wanted to use computers as tools. To take the naive user away from arcane commands and system dialogs, scientists delved deeper into cognitive science, which as a collective, included cognitive psychology, artificial intelligence, linguistics, cognitive anthropology, and the philosophy of mind. Researchers made pioneering efforts studying how people interacted with technology, even if they weren't quite sure what HCI initially was. They then redesigned software (and computers) to improve the computing experience, boosting productivity. Mouse technology and desktop icons made it easier for people to work with technology. Even though HCI was originally born in the academic home of computer science, and its primary function was bettering personal productivity applications, such as the software mentioned above (text editors

and spreadsheets) the field became increasingly diverse." It quickly expanded to encompass visualization, information systems, collaborative systems, the system development process, and many areas of design. HCI is taught now in many departments/faculties that address information technology, including psychology, design, communication studies, cognitive science, information science, science and technology studies, geographical sciences, management information systems, and industrial, manufacturing, and systems engineering." (Carroll, J. 2013) What Carroll tells us here is that HCI has become less focused on just a single aspect of its original core concepts and methods and it is now much broader than computer science itself. HCI in the 1980s began to step away from the individual and generic cognitive behaviour into three different ways. First was to have more of an ethnographic approach to design. This ethnographic approach includes social and organizational computing, accessibility for the elderly, the cognitively and physically impaired, and for all people, and for the widest possible spectrum of human experiences and activities. Ethnography was seen by some as a fashionable term but others believed that ethnography was something more. " It is less about method and more about the lens through which human activities are viewed." (Blomberg et al. 2003). The principle of ethnography Blomberg is trying to explain here is that if you are trying to understand something you know little about, you must encounter it firsthand. In the 1990s, ethnographically informed design had its interest amplified once more thanks the arrival of the internet, accelerating " the move of information technologies out of the workplace and into the homes" (Blomberg et al. 2003), and other non-work environments. This caused the

second wave for HCI to move away from the individual desktop, with the expansion upon " desktop office applications to include games, learning and education, commerce, health and medical applications, emergency planning and response, and systems to support collaboration and community" (Carroll, J. 2013). These systems to support communication, such as e-mail, brought about an influx of interest in the possibilities of these communication channels as people were not just interacting with computers but they were able to interact with each other through the computers. Collaborative activity saw the rise of several tools we have today, for example, wikis, blogs, social networking, and social bookmarking, just to name a few. These new mechanisms also created online auctions, soft sensors and crowd sourcing. However, today, this area of Human-computer interaction is now referred to as social computing (Carroll, J. 2013). And finally, the third was the expansion upon early graphical user interfaces so to include a multitude of " interaction techniques and devices, multi-modal interactions, tool support for model-based user interface specification, and a host of emerging ubiquitous, hand-held and context-aware interactions". (Carroll, J. 2013). Carroll mentions hand held devices, for example, the mobile phone which appeared in the mid1980s and even earlier than that, the appearance of the laptop. These devices created an explosion of diversification for designers and developers. Not to mention how these devices have led to the pervasive inclusion of computing into further human factors such as cars, homes, appliances and apparel. These three ways have created some of the fastest growing markets today and many believe even

though we have considerably distanced ourselves from the desktop, it is still very important.

## Chapter 3

### Gesture Recognition & its audience

What is the current state of GR in Live music performances? (In depth detail).

Who are the leaders in this field? Why is there a gap between the audience and the performer? / What is causing this problem? Gestural interaction with computational devices, and therefore the recognition of gestures, has been the focus of research throughout many areas of HCI since the development of early applications in the 1960's; such as Ivan Sutherland's Sketchpad (Sutherland, I. 1964) which used an early form of stroke-based gestures using a light pen to grab and manipulate graphical objects on a tablet display. Waren Teitelman was one of the first researchers to develop a trainable gesture recognizer that could classify hand drawn characters in real-time. Several other pen-based recognition systems followed in the 1960's and the 1970's, such as the GRAIL system (Ellis et al., 1969) or in the AMBIT/G system (Christensen, 1968); with this form of interaction now being widely accepted throughout the HCI community (Karam, 2006). Glove and magnetic sensor based systems, such as the work found in Bolt (1980), Zimmerman et al. (1986), Sturman et al. (1989), Quek (1994) and Wexelblat (1995), received a small amount of attention from researchers throughout the 1980s and early 1990s but this research was limited due to the large expense and technical requirements of the sensor technology. It was not until Freeman and Weissman (1995) first demonstrated a camera based system that enabled gestures to control the volume and channel function of

<https://assignbuster.com/can-gesture-recognition-in-a-live-digital-music-performance-media-essay/>



a television that the field of computer-vision rapidly started to grow. For the next decade, the HCI gesture recognition community could be generally segmented into two 4explicitly programming a machine to look for value a followed by value b and then value research categories, glove based recognition systems and computer-vision based recognition systems. With the breakthroughs and the maturity of modern technologies, as mentioned in the introduction, it is still extremely difficult for a performer to interact with a computer in the same way as they would interact with other performers live on stage.

## **Chapter 4**

## **Chapter 5**

### **Bridging the Gap.**

Can Gesture Recognition software shrink the gap? How do they engage the audience? Are the examples above successful?

## **Chapter 6**

## **Chapter 7**