Cinemacraft: Immersive Live Machinima as an Empathetic Musical Storytelling Platform

Siddharth Narayanan
Electrical and Computer Engineering
Virginia Tech
nsiddh3@vt.edu

Ivica Ico Bukvic
Institue for Creativity, Arts & Technology
Virginia Tech
ico@vt.edu

ABSTRACT

In the following paper we present Cinemacraft, a technology-mediated immersive machinima platform for collaborative performance and musical human-computer interaction. To achieve this, Cinemacraft innovates upon a reverse-engineered version of Minecraft, offering a unique collection of live machinima production tools and a newly introduced Kinect HD module that allows for embodied interaction, including posture, arm movement, facial expressions, and a lip syncing based on captured voice input. The result is a malleable and accessible sensory fusion platform capable of delivering compelling live immersive and empathetic musical storytelling that through the use of low fidelity avatars also successfully sidesteps the uncanny valley.

1. INTRODUCTION

Recent revolution in the area of off-the-shelf immersive technologies and phenomenology has changed the way users interact with games, media, and the arts. Creative projects have actively adopted the Microsoft Kinect [1], along with an array of affordable alternative all-in-one consumer-level motion capture devices to explore novel interactions and perceptions. Music-centric research, in particular, has focused on generating sound through user's movement [2], a dancer being the ideal user candidate due to the high level of body awareness, physical ability, and movement vocabulary. Experiments have also integrated human interaction into performance, either as stylized body movements [3] or through the use of virtual interfaces [4]. Some have gone as far as to create new environments for musical composition and performance [5]. Video games have also served as a rich foundation for artistic expression using immersive devices through machinima [6] and digital puppetry [7].

These new modes of interaction are also significant boon to the level of immersion in video game technology. Studies have shown that an embodied interaction is correlated to players' increased engagement [8] and a stronger affective experience [9]. Most modern video games, however, continue to be played using low fidelity interaction devices–interactions between video game characters and the user are often accomplished through keyboard presses or joystick manipulation for various body motions, ranging from simple actions such as walking or jumping to more elaborate tasks like opening doors or pulling levers. Such an approach can often result in non-natural and potentially limiting interactions. Such interactions also lead to profoundly different bodily experiences for the user and may detract from the sense of immersion [10]. Modern day game avatars are also loaded with body, posture, and animation details in an unending quest for realism and compelling human representations that are often stylized and limited due to the aforesaid limited forms of interaction. These challenges lead to the uncanny valley problem that has been explored through the human likeness of digitally created faces [11], and the effects of varying degrees of realism in visualizing players' hands [12]. Some titles have taken an alternative approach towards sidestepping the uncanny valley by utilizing simplified cartoon-like graphics. For instance, in Minecraft, a sandbox-like gaming environment that has attained an unprecedented level of popularity, the simple avatar movements and highly stylized cartoon- and 8bit-like graphics [13] make it "work" because of a strong mix of the game's aesthetic sensibility, open-ended design, mechanics, development history, and the creative activities of its players [14]. A growing number of research projects explore Minecraft as a platform for enhanced immersion [15] and with full body and facial immersion [16]. However, with the exception of HTC Vive's VR implementation of Minecraft [17] they have remained by and large confined to complex, difficult to reproduce setups. In addition, such projects often require motion-capture data to be aligned with other information, resulting in a complicated endeavour when utilizing a combination of devices [18]. This limitation also impedes on the level of immersion the platform is capable of delivering. Mutli-sensory immersive virtual environments have been shown to be capable of inducing emotional responses [19] and immersion has also been shown to significantly depend on the aspects of virtual presence like emotion [20]. Such emotional immersion [21] also promotes the building of a bond between the user and in game avatars in the virtual world. In particular, it helps in building empathy for the characters, an important aspect of storytelling [22] and also a significant challenge for the video game industry, where delivering compelling characters continues to be one of the primary challenges [23].

While previous related work has focused on promoting empathy, primarily through musical experiences [24], visual modality is also seen to contribute significantly to the perception of emotion in a musical performance. In par-

ticular, the ability to experience expressive body language has been shown to be important in perceiving emotional intensity of a performance [25]. Although typically unintended and similar to the expressive gestures that accompany speech [26], these ancillary gestures serve to communicate information which clarifies performer's emotional intentions. It is through these ancillary gestures that people are able to differentiate between musicians playing in a restrained and expressive manner even in the absence of audio [27]. Expressive body movements have also been shown to be significant for musicians in conveying emotional intentionality as a visual component during a musical performance to enhance the observer's identification with the performer [28]. This has also been explored through the integration of motion-capture and video-analysis techniques into music-centric research [29]. However, the use of consumer-level motion capture devices for the immersive, embodied, and affective computing, as well as compelling story telling within the context of performative art has yet to be fully explored.

1.1 Motivation

It appears the potential for immersive environments to serve as a compelling and easily accessible platform for musical expression, empathy, and storytelling remains largely underutilized. This is in good part due to the aforesaid uncanny valley challenge, as well as due to commonly complex, site-specific, and/or cost-prohibitive nature of the current solutions. Inspired by the successes of projects like Minecraft, we see this as an opportunity to introduce an accessible and affordable alternative that seeks to sidestep uncanny valley by employing cartoon-like environment without sacrificing the emotional and expressive depth. As evidenced by the cartoon and gaming industry, such a platform would not have to sacrifice the empathy or the storytelling potential. In addition, as part of this paper we see the newfound white space as an opportunity to broaden the definition of interfaces for musical expression, to incorporate enabling technologies that can project a compelling musical experience in an immersive technology-mediated environment, including telematics. In this paper we focus primarily on the machinima-like implementations that leverage existing gaming environments as a potentially compelling storytelling platform.

2. CINEMACRAFT

Cinemacraft is a novel technology-mediated immersive machinima platform for collaborative performance and musical human-computer interaction. It innovates on a custom, reverse-engineered version of Minecraft to offer a collection of live theatrical and cinematic production tools, and leverages the Microsoft Kinect HD for embodied interaction, including posture, arm movement, facial expressions, and through the sensory fusion lip syncing based on captured voice input. It is designed as an out-of-box turnkey solution that sidesteps the uncanny valley by utilizing the extremely popular sandbox-like gaming environment, for simple yet compelling storytelling along with multiple live camera views, scene changes, subtitles, lip sync, production-centric stage cues, and virtual audience. The platform is aimed at extending the frontiers of collaborative content creation as well as broadening audience impact to enhance creativity and emotional experiences.

2.1 The OPERAcraft Lineage

The OPERAcraft platform [30], a precursor to Cinemacraft was envisioned as an environment to aid creativity and thinking skills and better self-expression, with particular focus on the K-12 education opportunities.

Figure 1. Inspiring K-12 students to create stories through a live production of Operacraft

It was built as an arts+technology education platform where students could write a story and libretto, build a virtual set, costumes or virtual character skins, and ultimately control the characters within the virtual setting in a live performance accompanied by live singers and musicians. Many of these affordances are inherent to Minecraft platform– users can easily sculpt the landscape, interact with it, and change their own appearance. Others were added as part of the reverse engineering effort, resulting in a mod that is deeply integrated into Minecraft's core. These include character lip syncing based on the singer's input processed through the Pd-L2Ork [31] and forwarded to a FUDI-based parser via a UDP socket embedded inside reverse-engineered version of Minecraft, audience subtitles and stage cues only visible to the actors, ability to change between discrete arm positions and interpolate between them to provide rudimentary body language, and near-instantaneous scene changes through coordinated character teleportations and scene cross-fades. In an ongoing pursuit of building a compelling real-time machinima production platform, the second generation of OPERAcraft introduced in the fall 2015 as part of the second opera production offers additional affordances, including multiple camera views and cameras that are only visible to the actors, invisible bystanders, as well as stability improvements and optimizations that allowed the mod to scale beyond the original limit of five actors.

2.2 New Platform

Cinemacraft builds on the live performance capture aspects of the existing system through the integration of the Microsoft Kinect HD [1] device to provide a more immersive and expressive embodied experience in a virtual world including both kinematic data and facial expressions. It

in many ways supplants previously keyboard controlled arm expressions and instead provides full body immersion to the extent allowed by the simple skeletal structure of the Minecraft avatars that lack hands, elbows, and knees, and further enhances expressivity by tracking facial expressions. Based on the user tests and audience feedback, the avatar remains compelling despite the minimal character design due to the reported feeling of sentience offered by the user's real-life body motion and facial expressions. As a result, the avatar can show a dynamic range of emotional reactions and responses. In cinematic terms, the avatar no longer appears to be merely acting. Rather, it is the actor who is responding to their projection in and the situational awareness of the virtual environment. Spontaneous reactions like squinting against a sudden bright light help to humanize characters and make them more compelling than current game characters that seem shallow and with whom we have a hard time forming compelling, coherent relationships [23].

Figure 2. Cinemacraft - Avatar in action

Sensory fusion and multi-sensory stimulation has been shown to help in illusorily experiencing a virtual body and virtual body parts [32]. Its use within the context of embodied interaction can be particularly useful in situations where one sensory input does not provide adequate resolution and could benefit from secondary inputs that improve its accuracy and fidelity. In Cinemacraft, such a situation can be observed when trying to monitor mouth movement. Kinect HD, particularly when a user is located farther away from the sensor, so as to allow for full-body capture, is not capable of providing accurate capture of fast and varied shapes generated by the mouth. In this case, sensory fusion through the use of voice analysis can be an appropriate way of complementing visual capture. Similarly, in situations where a user has their mouth opened and are making no sound, visual capture needs to take precedence as the only reliable source of information. The current Cinemacraft version implements OPERAcraft-based avatar lip syncing based on the simple transient and melisma detection that is cross-pollinated with the Kinect HD-based capture with the two swapping precedence based on the context.

2.3 Modes of Interaction

Cinemacraft offers different modes of embodied interaction captured by Kinect HD, namely regular, mirrored, and upper torso only. For instance, if the user prefers to both control the avatar and act out the gestures and facial expressions they can gesticulate to the other players to complement their speech and chat messages and thereby increase the effectiveness of the conversations, while still being able to navigate the expansive landscape outside the range afforded by the area monitored by Kinect HD using more conventional controls (e.g. keyboard). The same can be also applied in the gaming scenarios, as well as hybrid situations where a separate user controls the avatar while an operatic singer, for instance, provides only upper body language. Similarly, the mirroring mode has been added to explore illusory experience interactions with the avatar, most notably through the Mirrorworlds research project focusing on the study of integration of physical and virtual mirrored presence [33]. These modes provide an opportunity to draw parity between the different approaches to machinima and also open new exciting possibilites for sensory fusion, with the introduction of HMD displays along gesture-based and haptic controllers. Cinemacraft also inherits a battery of OPERAcraft's cinematic tools, empowering user to explore the methods of live machinima production, including live theatrical play, as well as produced cinema. The virtual audience feature, that enables audience members to freely roam the scene, or the ensuing world in which the storytelling takes place, offers new research opportunities in the study of perception of story telling, drama, and empathy as a function of vantage point.

3. IMPLEMENTATION

Cinemacraft mirrors the user's motions and expressions in real time within Minecraft using the Kinect HD sensor and a custom C# client that leverages Kinect HD API. Our platform expands on the previously modified Minecraft 1.5.2 codebase from the OPERAraft project to make acting more expressive and closer to the levels of post-produced machinima. The C# Kinect application and Cinemacraft mod for Minecraft are packaged as a drop-in mod and an independent executable file, respectively. A major consideration for our setup was accessibility. For this reason, although earlier prototypes relied on two first generation Kinects, due to their observed inability to simultaneously do body and face tracking, the current implementation relies on just one second generation Kinect HD device responsible both for kinematic and facial tracking.

In order for the Kinect HD client to interface with the Minecraft mod, core Minecraft was reverse engineered and retrofitted with a FUDI-compliant protocol [34] capable of parsing remote messages. Its simpler version is already found in OPERAcraft where it was used to coordinate various aspects of the production, including switching camera angles, lip syncing as detected by the singers' microphones, subtitles, and stage cues. As a result these could be handled remotely through multiple distributed Pd-L2Ork clients [31].

One of the main objectives of the system is to provide a compelling emotional delivery of storytelling within the context of arts. Therefore, it was important to test how re-

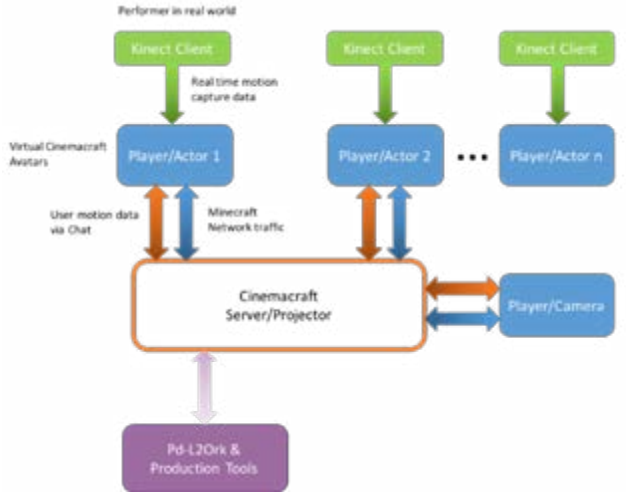

Figure 3. Cinemacraft - Kinect, client, server architecture

alistically the avatar could reflect that user's kinematic and facial expression. As a result, the system was developed using iterative design approach. Its testing focused primarily on assessing the perceived visual fidelity of mirroring user's interaction.

3.1 Sensory Fusion

The emphasis on ease of use and reliance only a single Kinect HD device requires our implementation to essentially stretch the limits of the current Kinect HD API. Despite its improved resolution over the first generation, Kinect HD is still best suited for face tracking in close proximity which limits its ability to track body. In turn, our implementation offers accurate simultaneous full body and facial tracking. Further still, we have identified problems with Kinect's machine learned library of postures and facial expressions that have resulted in a prevalent number of false positives pertaining to eye winks, eyebrow movement, and eyeglass detection. For this reason in the current iteration eye winks and eyebrow movement has been temporarily disabled. We aim to address these challenges through an additional layer of sensory fusion. One possible approach is to utilize a low-latency computer vision algorithm on the area of the image that Kinect HD identifies as containing the tracked human face. Another could be further enhancing face detection with infrared video feed inherent to Kinect HD and look for high reflectivity of eye pupils.

We have envisioned a platform with parallel pipelines of Audio Inputs, Kinect API and Computer Vision optimization and learning for improving facial Expressions, with all 3 working together to further refine the platforms capabilities through sensory fusion. To address Kinect HD's limited ability to detect mouth shapes, we merged the depth camera data with the captured audio input. Here, the sensory fusion allowed us to use voice detection to combine the performer's audio with the facial tracking data and thereby improve detection of minor gestures and expressions which may not be otherwise captured due the technical limitations of the two distinct approaches to monitoring user's input. For instance, doing so enabled us to animate mouth motion through captured audio that exceeds the resolution of 30 frames per second, as well as audio-centric outliers, such as the cartoon-like quivering of lips in a sung operatic melisma. Conversely, we also relied on the depth camera feed in situations when no sound was being made, e.g. to portray gaping mouth. The ensuing implementation utilizes a simple logic by which the two sensory inputs are given precedence.

4. REAL WORLD TESTING & USER FEEDBACK

To date, Cinemacraft was showcased as part of two high profile exhibitions with two additional opportunities pending. The team has used such opportunities to iteratively improve upon and refine the design, as informed by the outcomes demonstrations and real world user feedback. In particular, the prototype was showcased at Virginia Tech's official exhibit at South by Southwest 2016 [35], and as part of ICAT day showcase at the Moss Arts Center in Virginia Tech [36]. More recently, Cinemacraft has also been chosen to be displayed at the Science Museum of Southwest Virginia [37]. As a result of the strong response to and interest in the tool, it has also been selected to be integrated in the Virginia Tech Visitor Center. Both exhibits are scheduled to open in the winter of 2017.

User feedback from real world assessment also helped us iteratively improve the platform performance. A number of problems were eliminated during the early assessment, a prominent one being the jittery avatar movements due to the discrete and rapidly fluctuating values from the real time motion capture data from the Kinect HD. This was solved by discriminating against implied joint positions and data points suggesting sudden jumps. Further issues included the structure of arms and legs rotations. The most notable user feedback from the tests was about their ability to understand avatar's inference of gestures that had a strong reliance on elbows despite the lack of such a joint in the avatar's skeletal structure. Given the lack of elbows, hands, and knees, the body anatomy of a Minecraft characters was significantly different to the human form, due to which the position and nature of rotation of joints had to be modified. This is one particular research area we are keen on exploring further. Users also expressed the desire for wider range of actions which would necessarily require a fundamental re-thinking of how human actions can be translated within this context. Another significant feedback was the importance of facial expressions and that the real-time tracked facial movements massively improved emotive qualities of the avatar. Tracking and representation of eye winks and eyebrow movements were disabled for certain periods of the assessment to reduce the number of false positives, and this seemed to sharply reduce the users' level of immersion. There was also strong positive feedback for the ease of stepping in and out of the virtual world. The application has been built to dynamically lock onto the closest user in the scene and is robust in handling their movements when they leave and walking back into the designated space.

4.1 Minetest

Throughout Cinemacraft's development we have interacted with a number of machinima studios worldwide to gain

their feedback on the mod and the interface. One of the most significant observations coming out of this interaction was the recently proposed switch from Minecraft to an open source alternative–Minetest [38]. Minetest offers several potential advantages over Minecraft. The Minecraft code is not readily modifiable and requires a community-driven effort to decompile JAVA runtime into a human-readable API. As such its forward compatibility is at best cumbersome. The legal implications of modding Minecraft are also not entirely clear, suggesting Microsoft by default owns all modded code, and as a result the distribution of the ensuing deeply integrated mod is difficult if not impossible. Minetest on the other hand, is strikingly similar yet highly moddable voxel game engine that covers a majority of Minecraft features. The in-game client server interactions are also better handled in Minetest with modded servers sending textures and other required resources to the clients, unlike Minecraft that may require resources to be downloaded separately. As a result, we are currently in the process of shifting mod onto the Minetest platform.

5. CONCLUSION

We have designed a new platform for immersive performance-centric interaction for the purpose of musical expression and story telling. The new platform is inspired by the success of Minecraft and builds on its approach that successfully sidesteps the uncanny valley. Our results so far are promising. As evidenced by audience and user feedback, including the machinima community, as well as interest in the platform, we were able to create a high level of immersion by combining multiple interaction techniques into a single system despite relying on a cartoon-like low fidelity environment. Extending sophisticated technology like immersive VR and gesture tracking to easy markerless motion capture our performers could control their avatar with relative ease and accuracy without extended training sessions.

6. BROADER IMPACT

Cinemacraft offers opportunities to extend virtual presence and consequently outreach by allowing audiences to engage with the production directly in-game. The team envisions the ensuing implementation being appropriate in a broad range of live and post production scenarios, beyond its original intent, from machinima movie-making to theatre. Studies have shown that learning movie and theatre production skills help to instill the sense of ownership, confidence and self-belief in students [39]. Minecraft has already been effectively used as an education tool through the successful MinecraftEdu [40]. Our shift towards the Minetest platform and an increased reliance on sensory fusion has a potential improve such an educational experience as a live learning tool.

7. FUTURE WORK

The core functionality of Cinemacraft has been developed keeping in mind, future opportunities for exciting improvements that could be implemented in the future. Cinemacraft's shift to Minetest has been made to support versatile additions. In particular is the support for multiplayer interactions which would involve a multitude of body motions, gestures and expressions. Such multi-user interactions have been shown to be very engaging and enrich immersion by giving the interactions with virtual environments an interpersonal dimension [41]. Open questions also remain around the use of Kinect HD. Current reliance on Kinect HD hardware restricts the user to a relatively small physical area, and mapping movements around in a large virtual world is challenging. In addition, there also seem to be some inherent gaps in Kinect detection. One prominent instance of this was during the detection of the user blinks when the user tilted his head toward one side beyond a certain angle. Further research is required into the exact technical limitations of the sensor and how much optimization can be achieved through software manipulation. Most importantly, the uncertainty surrounding Kinect HD's future has prompted us to explore alternative open source solutions.

8. REFERENCES

[1] M. K. HD. (2017) Kinect Kinect HD. [Online]. Available: http://www.xbox.com/en-US/xbox-one/accessories/kinect

[2] T. Berg, D. Chattopadhyay, M. Schedel, and T. Vallier, "Interactive music: Human motion initiated music generation using skeletal tracking by Kinect."

[3] J. Ratcliffe, "Hand motion-controlled audio mixing interface," *Proc. of New Interfaces for Musical Expression (NIME) 2014*, pp. 136–139, 2014.

[4] T. Ahmaniemi, "Gesture Controlled Virtual Instrument with Dynamic Vibrotactile Feedback." in *NIME*, 2010, pp. 485–488.

[5] C. Honigman, A. Walton, and A. Kapur, "The third room: A 3d virtual music paradigm," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2013, pp. 29–34.

[6] R. Kastelein. (2013) The Rise of Machinima, the Artform. [Online]. Available: http://insights.wired.com/profiles/blogs/the-rise-of-machinima-the-artform

[7] E. Polyak, "Virtual Impersonation Using Interactive Glove Puppets," in *SIGGRAPH Asia 2012 Posters*, ser. SA '12. New York, NY, USA: ACM, 2012, pp. 31:1–31:1. [Online]. Available: http://doi.acm.org/10.1145/2407156.2407191

[8] N. Bianchi-Berthouze, "Understanding the role of body movement in player engagement," *Human–Computer Interaction*, vol. 28, no. 1, pp. 40–75, 2013.

[9] N. Bianchi-Berthouze, W. W. Kim, and D. Patel, "Does body movement engage you more in digital game play? And Why?" in *International Conference on Affective Computing and Intelligent Interaction*. Springer, 2007, pp. 102–113.

[10] J. Parker, "Buttons, simplicity, and natural interfaces."

[11] S. Lay, N. Brace, G. Pike, and F. Pollick, "Circling Around the Uncanny Valley: Design Principles for Research Into the Relation Between Human Likeness and Eeriness," *i-Perception*, vol. 7, no. 6, p. 2041669516681309, 2016. [Online]. Available: http://dx.doi.org/10.1177/2041669516681309

[12] F. Argelaguet, L. Hoyet, M. Trico, and A. Lécuyer, "The role of interaction in virtual embodiment: Effects of the virtual hand representation," in *Virtual Reality (VR), 2016 IEEE*. IEEE, 2016, pp. 3–10.

[13] N. Garrelts, *Understanding Minecraft: Essays on Play, Community and Possibilities*. McFarland, 2014.

[14] S. C. Duncan, "Minecraft, beyond construction and survival," *Well Played: a journal on video games, value and meaning*, vol. 1, no. 1, pp. 1–22, 2011.

[15] S. Choney. (2016) Microsoft Stores offering free Minecraft VR demos on Oculus Rift. [Online]. Available: https://blogs.microsoft.com/firehose/2016/09/29/microsoft-stores-offering-free-minecraft-vr][][http://dl.acm.org/citation.cfm?id=2910632

[16] N. Viniconis. (2011) MineCraft + Kinect : Building Worlds! [Online]. Available: http://www.orderofevents.com/MineCraft/KinectInfo.htm

[17] Vivecraft. (2016) Vivecraft. [Online]. Available: http://www.vivecraft.org/

[18] B. Carey and B. Ulas, "VR'SPACE OPERA': Mimetic Spectralism in an Immersive Starlight Audification System," *arXiv preprint arXiv:1611.03081*, 2016.

[19] J. Bailey, J. N. Bailenson, A. S. Won, J. Flora, and K. C. Armel, "Presence and memory: immersive virtual reality effects on cued recall." Citeseer.

[20] M. Lombard and T. Ditton, "At the heart of it all: The concept of presence," *Journal of Computer-Mediated Communication*, vol. 3, no. 2, pp. 0–0, 1997.

[21] J.-N. Thon, "Immersion revisited: on the value of a contested concept," *Extending Experiences-Structure, analysis and design of computer game player experience*, pp. 29–43, 2008.

[22] M. Simon and S. Mayr, *Immersion in a Virtual World: Interactive Drama and Affective Sciences: Interactive Drama and Affective Sciences*. Anchor Academic Publishing, 2014.

[23] C. Buecheler. (2010) Character: The Next Great Gaming Frontier? [Online]. Available: http://www.chicagotribune.com/sns-gamereview-feature-characters-story.html

[24] F. Visi, R. Schramm, and E. R. Miranda, "Use of Body Motion to Enhance Traditional Musical Instruments." in *NIME*, 2014, pp. 601–604.

[25] H. Blyth, "Cross-Modal Perception of Emotion from Dynamic Point-Light Displays of Musical Performance," 2015.

[26] M. M. Wanderley, "Quantitative analysis of non-obvious performer gestures," in *International Gesture Workshop*. Springer, 2001, pp. 241–253.

[27] J. W. Davidson, "Visual perception of performance manner in the movements of solo musicians," *Psychology of music*, vol. 21, no. 2, pp. 103–113, 1993.

[28] B. W. Vines, C. L. Krumhansl, M. M. Wanderley, and D. J. Levitin, "Cross-modal interactions in the perception of musical performance," *Cognition*, vol. 101, no. 1, pp. 80–113, 2006.

[29] V. Bergeron and D. M. Lopes, "Hearing and seeing musical expression," *Philosophy and Phenomenological Research*, vol. 78, no. 1, pp. 1–16, 2009.

[30] I. I. Bukvic, C. Cahoon, A. Wyatt, T. Cowden, and K. Dredger, "OPERAcraft: Blurring the Lines between Real and Virtual." in *ICMC*, 2014.

[31] I. Bukvic, "A Behind-the-Scenes Peek at World's First Linux-Based Laptop Orchestra–The Design of L2Ork Infrastructure and Lessons Learned."

[32] A. Lcuyer, "Playing with Senses in VR: Alternate Perceptions Combining Vision and Touch," *IEEE Computer Graphics and Applications*, vol. 37, no. 1, pp. 20–26, Jan 2017.

[33] N. F. Polys, B. Knapp, M. Bock, C. Lidwin, D. Webster, N. Waggoner, and I. Bukvic, "Fusality: an open framework for cross-platform mirror world installations," in *Proceedings of the 20th International Conference on 3D Web Technology*. ACM, 2015, pp. 171–179.

[34] FUDI. (2017) FUDI. [Online]. Available: https://en.wikipedia.org/wiki/FUDI

[35] V. Tech. (2016) South by Southwest 2016. [Online]. Available: http://www.vt.edu/sxsw.html

[36] A. Institute for Creativity and Technology. (2016) ICAT Day 2016. [Online]. Available: https://www.icat.vt.edu/content/icat-day-2016

[37] V. Tech. (2017) Science Museum of Western Virginia. [Online]. Available: http://www.icat.vt.edu/smwv

[38] Minetest. (2016) Meet Minetest. [Online]. Available: http://www.minetest.net/

[39] A. Swainston, N. Jeanneret *et al.*, "Wot Opera: A joyful, creative and immersive experience," in *Music: Educating for life. ASME XXth National Conference Proceedings*. Australian Society for Music Education, 2015, p. 99.

[40] Microsoft. (2016) Minecraft Education Edition. [Online]. Available: https://education.minecraft.net/

[41] M. Loviska, O. Krause, H. A. Engelbrecht, J. B. Nel, G. Schiele, A. Burger, S. Schmeißer, C. Cichiwskyj, L. Calvet, C. Griwodz *et al.*, "Immersed gaming in Minecraft," in *Proceedings of the 7th International Conference on Multimedia Systems*. ACM, 2016, p. 32.