Beyond the Audience Seat: The recording and production of immersive opera and interactive concerto programmes for VR experiences

Hans-Peter Gasselseder
Aalborg University
Dept. of Communication & Psychology
Musikkens Plads 1
9000 Aalborg, Denmark
maestroso@gmail.com

Maria Kallionpää
Hong Kong Baptist University
Rm AST801, Sing Tao Building
Hong Kong Baptist University
Kowloon Tong, Kowloon, Hong Kong
makallio@hkbu.edu.hk

New modes of expression have always defined a central attribute within the ambitions of classical music. The extended realities (XR) hold the potential to expand this expressive repertoire by transporting audiences into the middle of the action within an immersive music performance. We argue that implementing these innovative formats in contemporary opera performance, composition, production, and recording opens up new aesthetical paradigms that can reciprocally inform the artistic disciplines and reach a wider demographic of audiences. To our knowledge, this paper presents the first ever interactive virtual reality (VR) production of an entire opera, "She" (2017) and "Croak" (2018). The recordings devise 3-dimensional 360-degree high-resolution video and 3rd-order Ambisonics audio from multiple perspectives within a bespoke software interface. Following a discussion of both operas from the composer's and engineer's perspective, the paper will present an enhanced rendition of "CroakVR", a 6-degrees-of-freedom (6DoF) volumetric simulation of its corresponding opera, that raises the bar of interactivity by enabling audiences to freely explore various mimetic and diegetic perspectives of its staging. Furthermore, we discuss on-going work on "Climb!VR", a 6DoF volumetric VR production devising multiple Ambisonics recordings of an interactive electro-acoustic piano concerto "Climb!" based on ludonarrative structures. It is argued that VR recordings can amend the life span of musical compositions by offering listeners a concert experience that can extend beyond the walls of the physical concert hall.

Opera. Virtual reality. 6DoF. Volumetric. Ambisonics. Recording.

1. INTRODUCTION

Opera requires a constant evolution in its design to stay interesting for audience and creatives. We argue that implementing innovative technologies in contemporary opera performance, composition, production, and recording opens up new aesthetical paradigms that can also reciprocally inform the artistic disciplines and reach a wider demographic of audiences.

To our knowledge, this paper presents the first ever interactive virtual reality production of two entire operas, "She" (2017) and "Croak" (2018). Both productions devise 3-dimensional 360- degree high-resolution video and 3rd-order Ambisonics audio from multiple perspectives within a bespoke software interface. Following a discussion of both operas from the composer's and engineer's

perspective, the paper will present on-going work on "Climb!VR", a 6-degrees-of-freedom volumetric VR production devising multiple Ambisonics recordings. The experience is simultaneously an interactive, virtuoso piano composition for a professional pianist and a computer game that also uses a specifically designed smartphone application for audience members. Using these case studies, the paper examines methods to facilitate the production of interactive volumetric simulations as well as its implementation within a three-dimensional spatial framework. In addition to laying out the design, recording, and post-production guidelines for both the experiential gains of naturalistic simulation as well as creative manipulation, we discuss how artefacts arising from 6DoF post-conversion may be mitigated for perceptual realism, as well as enhanced and highlighted for creating a hybrid format that

embraces both the stage presence as well as the artificial nature of a concert setting. This will be weighed against the circumstances of the recording situation where limitations set by the content, location, staging, budget, and compliance of artistic team members contribute hierarchically to the selection, layout and handling of recording equipment as well as post- production workflows. Likewise, we showcase the need to consider for diegesis, agency, and attention centring of the visual and auditory field at the stage of the recording plan.

> Ultimately, the presented case studies aim to evolve the experience of a 'stage presence' for the audience in synergy with the stage production to a degree that the newly united form stands as an artwork in its own right.

2. ARTIFICIAL AND VIRTUAL REALITIES IN CLASSICAL MUSIC

Classical music composers have always been eager to find new modes of expression, as well as techniques and technologies that would enable them to realise their ambitious artistic visions. Moreover, the possibilities of the artificial and virtual reality tools have provided efficient solutions for creating more immersive music recordings, which may help to raise interest in new audiences that would not normally pay attention to classical music (Kallionpää, Chamberlain and Gasselseder 2018). Such technologies can help to keep the classical music culture alive and evolving, instead of becoming a museum for a diminishing group of people who actively go to concerts. Because contemporary classical music strongly leans on live performances, such compositions are often only known by the listeners who attended their performances. The authors of this paper argue that VR recording techniques can amend the life span of musical compositions by providing the listeners an actual concert experience beyond the walls of the physical concert hall.

The design process of different interfaces and software also affects the artistic work of the composers: innovative use of technologies unlocks a vast potential of new kind of musical forms, sound production techniques, and notation. Santini´s (2018) research on live-generated notational interfaces is an illustrative example of the use of augmented realities in classical music composition. The LINEAR system is an AR environment developed for the purposes of musical performance: with the help of an AR iPhone application, the performer can create virtual objects that are "rendered in real-time and superimposed to the real environment" (Santini 2018). The system is used as an electronic musical interface, as well as a (live) graphical notational tool. As discussed by

the author of the system, it is especially useful when the composer needs to communicate how to play specific extended techniques, as it may be difficult to express the exact physical movements necessary to obtain the desired sonic outcome with the means of traditional notation. Instead of having to use complex explanations on the musical score, the virtual objects directly show the performers how to access them. Furthermore, the presence of various perspectives and interaction with the virtual bodies in LINEAR enable a higher level of flexibility than what would be normally achievable (Ibid).

Composer and film director Michel van der Aa attempts to engage his audiences in immersive operatic experiences. In his works, he has applied a variety of innovative technologies that, instead of functioning as show elements, form a vital part of the dramatic structure. Van der Aa´s opera "Blank Out" (2013) utilises 3D projection techniques in combination with live-performance, pre-recorded material, and sound processing. He also explored with them in his film opera "Sunken Garden" (2011–2012). Moreover, the mixed-reality project "Eight" is an outcome of Aa´s collaboration with the virtual reality company The Virtual Dutch Men and the designer Theun Mosk, the purpose of which was to "create a unique, ground breaking fusion of musical theatre, VR, and visual art" (van der AA 2018). "Eight" is a musical installation, through which the audience members move one by one, equipped with VR glasses and headphones. Like in his other works, the role of the technology is to support the storytelling: the listeners manipulate physical and virtual objects and meet the main protagonist in her different life phases (Ibid).

Moreover, various augmented reality solutions have been developed for the needs of performing musicians. For example, Pianolens (An Augmented Reality Interface for Piano Instruction) was created to help learning of new music and to speed up the rehearsing process. The system consists of an interactive, dynamic music score that visualises musical objects intuitively on the "music sheet", demonstrates correct timing and fingerings, and allows the musician to move flexibly between various segments of the score (Strasnick et al. 2016).

3. CASE STUDY: SHE (2017), CROAK (2018), AND CLIMB (2017)

3.1 The composer´s perspective

Despite of being rather different compositions, "She", "Croak", and "Climb" were largely written simultaneously, which is why certain overlapping of the musical material can be detected (for example, the composer used similar material at the

"Summit"-part of "Climb" and the scene IV of "Croak"). However, the instrumentation, performing set ups, and the compositional concepts varied between each musical work, which is why each of them required an individual recording plan. A lot depended on the stage directors: whereas Luca Bargagna, the director of "She", opted for a relatively distanced, two-dimensional visual storytelling, the puppet theatre opera Croak´s director Anna Ivanova-Brashinskaya attempted to create a more interactive stage presence. On top of the stage design itself, also the human sized puppets used in the latter work set their own challenges: as both the puppets and the puppeteers were moving freely on the open stage, more three-dimensional recording approach was called for.

"Climb" drastically differs from the two above-mentioned works, as it is a relatively abstract composition (although it has a non-linear dramatic storyline). The latter work has taken its inspiration from the modern day dynamic computer game music systems: it is a composition for Yamaha Disklavier and interactive system, and, as discussed earlier, it also includes interactive visuals, a web archive, and a smartphone application (Kallionpää et al. 2017). Whereas both "She" and "Croak" are musical dramas that constantly engage a large number of musicians and actors on stage, "Climb" only requires one performer, which is why it was natural to attempt to create a VR environment from a more intimate perspective (from the perspective of the performer). Moreover, the recording plan of "Climb" was also intended to mimic a first-person perspective of a computer game, which was also the core of its compositional structure. To support the versatile artistic aims of each production, different camera- and microphone set-ups had to be designed.

3.2 Recording "She" and "Croak": the making of She360 and CroakVR

On top of their artistic purposes, the opera productions "She" and "Croak" showcased technical aspirations that distinguished them from the previous offerings in the genre. These circumstances were accounted for with regards to the chosen recording strategy. To the best of the authors' knowledge, these operas represent the first full-length recordings of an entire opera in an immersive format that supports 3-dimensional 360-degree video and audio for playback and control in virtual reality (VR; i.e. head movements as well as six degrees of freedom facilitated by volumetric simulations in "CROAK VR"). Apart from the idea of creating opera recordings for VR, a further goal was to maintain compatibility with legacy as much as next generation formats, with supported playback setups ranging from traditional channel-based systems (such as 5.1 surround sound) to full 360-degree object-based surround (either rendered binaurally or on an arbitrary number of loudspeaker pairs). The recording plan aimed to offer audiences spectator perspectives that would not be possible in a regular opera performance. However, the main difference in the selection of immersive recording techniques for both operas was due to circumstantial requirements set by the stage design and staging, musical material and orchestration, as well as legal considerations.

With regards to the video footage for "She", a two-dimensional 360-degree format was to account for the entirety of stage properties being simulated by projections on a translucent screen positioned anterior to the characters on stage. This positioning allowed spectators in VR to move around in the scene, looking at the stage in the front and at the orchestra at the back. Due to the play directors' intentional use of a stage design that was to mimic a two-dimensional perspective in reference to the mythological character of the narrative, it was decided to uphold this quality while extending its immersive potential by placing the main camera rig (Insta360 Pro in 2D at 8k resolution, see Insta360 2017) closer towards the characters on stage (approx. 2 meters to camera objects) as compared to standard setups in VR production (approx. 4-6 meters if shot close-up). This allows users to experience the characters from up close but also to zoom out of the equirectangular source (an option provided in the user interface of the bundled software) as to get a view of the overall 'canvas' that the stage design was intended to convey. For audio, the same rationale came to effect by placing a prototype 3rd-order Ambisonics microphone (Zylia ZM-1d; prototype kindly provided by the manufacturer; Zylia 2019) in front of the orchestra (i.e. at the level of the conductor). This acoustic perspective was selected to emphasise the staging (i.e. everything placed in front of the spectator), but also to add the sense of place (i.e. offering a more holistic acoustic representation of the situational context with less direct and more reverberant sound reaching the microphone) by positioning the microphone further back from the camera rather than placing both at the same position (which would represent the standard dictum in VR productions, see Schütze and Schütze 2018). In addition, the flexibility achieved by recording in Ambisonics allowed focusing on different sound sources during post-production as much as during real-time playback in the final product. To give an example of a typical use case, turning ones' head towards the magnetic resonator piano at the left side of the orchestra pit will turn its sound more prominent in comparison to the otherwise dominant orchestral texture within the mix. When reflecting on an early review of the methodological insights gained while recording "She", for the visual domain,

on the one hand, the added sense of 'being there' was achieved by moving closer and thus offering more detail on the character interactions. For audio, on the other hand, the counter-approach seemed most effective by moving further away from the scene and subsequently adding a holistic sense of space and situational context to the experience (see Gasselseder 2015 for some theoretical considerations related to this finding).

In contrast, the production of "Croak" exhibited a rather different set of criteria to be respected by the recording. Whereas the staging of "She" necessitated a two-dimensional approach with added depth (for the sake of supporting the sense of involvement as much as 'being there'), the production of "Croak" posed a challenge in the opposite direction. The production made little use of stage properties but rather had its characters move a lot and explore the depth of the stage to large extend. Moreover, the nature of the main characters being represented by life-size puppets controlled by puppeteers would have made spectators prone to confuse character relationships if the same two-dimensional recording setup had been used as in "She". Thus, rather than positioning the recording gear close-up (visual) / far-off (audio), the recording rationale opted for a wide and more distant view of the stage that embraced depth by means of a three-dimensional video (as compared to the two-dimensional approach applied in "She") and a more surrounding audio experience. This allows spectators to differentiate the suggested layers of diegesis that are inherent in puppet theatre where the representative characters ("puppets") are accompanied by puppeteers, and in the special case of opera, also by individual singers. Where "She" required depth in its depiction of the stage, "Croak" necessitated a sense of segmentation of the different hierarchical diegetic layers suggested by the positioning of individual actors/singers. With puppets commonly positioned in the mid-front, singers at the side edge of the stage, and puppeteers further behind the former two, a three-dimensional video in 360-degrees enables spectators to differentiate these (hierarchical) layers. Another complexity specific to "Croak" posed the stage lighting, which varied in colour, intensity and, in combination with the shadows cast by the puppets, caused interesting reflections onto the sides and rear of the concert hall. These reflections were mainly visible from the audience's perspective. In order to capture the depth of the stage as well as an alternative perspective of the hall from the audience's view, the main camera rig (Insta360 Pro in 3D at 6k resolution, see Insta360 2017) was placed approximately 6 meters at the edge between stage and orchestra pit. Furthermore, the primary rig was extended by a secondary 360-degree camera setup (2x Kodak

SP360 4k, see Kodak 2017) positioned between the first row of the audience seating area as well as above the conductor in the orchestra pit. For the final VR experience, this configuration enables a change of perspectives during playback when zooming out from the footage recorded on the primary camera and blending over to the material that was shot on the secondary cam-era rig. Due to the lack of available software allowing for the playback and control of different camera perspectives in VR, we developed a custom player accounting for the aforementioned requirements within the game development platform Unity. Opting for a 2.5D paradigm that involves the online-mapping of flat surfaces onto 3D-objects, the player software devises dynamic displacement maps derived from depth maps to extrude meshes and project the camera footage within a 3D-environment. The Z-depth streams were partially derived from the parallax of stereoscopic footage (discussed further below) and enhanced by adding custom drawn mappings to account for the hall geometry and introduce special effects at key points of the experience. In order to counteract overlap distortions arising from stitching multiple depth maps, rotoscoping of select characters and objects was deemed necessary for adding clarity as well as interactive features to be introduced in future versions of the player.

Within the player environment, the footage obtained from either camera perspective can be assigned to individual hotspots. These come to effect upon reaching a threshold value of the zoom control parameter and allow for seamless transitions and blending of both camera perspectives. In this way, if the user decides to switch perspective or focus on another object shown on stage or in the audience room she/he is always presented with the highest resolution of the chosen perspective. Moreover, as different perspectives are being blended between two pseudo-rooms of the same scene (i.e. stage and audience room), a careful consideration of camera positioning enables us to generate transitionary scenes within the virtual environment (i.e. extrapolating new virtual perspectives, not shot with physical camera positions, but generated from footage obtained from available camera perspectives) and approach a six-degrees-of-freedom (6DoF) volumetric simulation where users can move within the captured video footage. To explore this idea, an in-development version of "CROAK VR" uses a mixture of photogrammetry of a static scene while blending in extra layers that exclusively contain the moving parts of the scene (i.e. puppets and puppeteers, RGB+Depth data). As only the primary camera position was recorded in stereoscopic 3D, the required depth information (i.e. depth map using shades of black/white to represent distances from the camera) was

projected from the primary onto the secondary camera rig. However, this procedure involves a lengthy process of rotoscoping and shading moving objects onto the corresponding depth map of a scene. By the same token, the extrapolation of virtual perspectives may suffer of artefacts related to the lack of resolution and spatial synchronization between depth maps obtained from the different camera models and positions. Moreover, due to the nature of the flat capturing inherent to video, we are missing out on information that lies behind those objects facing the camera (i.e. accuracy of textures is two-dimensional but displaced within a three-dimensional volumetric simulation). Whereas this working method would pose the risk of artefacts separating acting characters from the remainder of the scene (as in the lack of accounting for shadows and blending silhouettes), the specific staging situation of Croak allowed making such sacrifices at the benefit of the overall experience. In this connection, it was found that in comparison to the unprocessed footage, the artefacts resulting from unsynchronised depth processing contributed to an appearance of puppet characters that was more abstract and divorced from reality, whereas the remaining acting parties (puppeteers, singers, orchestra/conductor, audience) were almost left untouched by the alterations. It is believed that this technique supports the user in distinguishing the different layers of the diegesis, which suggests three acting parties (puppet, puppeteer, singer) as an agent of a single character and thus has to tackle the issue of assigning agency of various expressions (appearance, action, singing) to a single originator.

As for audio, the staging as well as setting of "Croak" implied moving closer with the Ambisonics microphone (Zylia ZM-1d prototype) towards the stage action. Therefore, the microphone was placed between the stage and the orchestra (rather than in front of the orchestra as in "She") and also functioned as the centre of a Decca tree setup that was supported by an additional 12 spot microphones. The reasoning for a closer acoustic perspective for "Croak" was two-fold. First, the lack of a contractual agreement between the producers and the orchestra meant that no orchestral material could be released on a public recording. Thus, a solution had to be found that would allow separating the voices of the singers and choir on stage from the orchestra, which would later be replaced by realistic virtual orchestrations. For that purpose, a drier (i.e. less reverberant) signal was preferred from which voices could be isolated during post-production. This was made possible by a bespoke procedure that focuses individual sound sources in a first step (with the help of directional focusing achieved by Ambisonics recording techniques; as realised in software such as Zylia Studio Pro, see Zylia, 2019) and separate these

from the remaining interfering signals during a second step of post-production (i.e. using spectral editing as well as specialised software such as Audionamix ADX Trax Pro 3, Audionamix, 2019). Subsequently, a multitude of isolated signals derived from the virtual microphones of the Ambisonics signal is used to phase-cancel harmonic noise and sum the target sound source (i.e. the voice of a soprano during a loud orchestra tutti). Being able to isolate individual voices and instruments from a single physical microphone position allows for elaborate post-production techniques that facilitate mixing in object-based audio formats and thus enhance quality and flexible adaptation of immersive audio con-tent for different output formats. The second rationale for a closer acoustic perspective goes in line with the previously outlined reasoning behind the camera setup. The recording intended to capture the singers more up-close as to enable spectators to locate the source of a voice and identify characters at a higher accuracy, thus facilitating diegetic orientation. As a result, the audio tracks of the VR presentation place spectators in the middle of the action, with the singers and choir positioned at the front and the orchestra at the back of a 360-degree sphere.

Conversely to the findings on "She", the sense of spatial awareness within the audio recording of "Croak" was less invested into the room/environmental aspects of the concert situation rather than the dramaturgic nature of character relationships and its construction of absorption (see Gasselseder 2015). In contrast, the video material served as a means of enhancing the sense of space and context, a function that in "She" was predominantly covered by the audio recording. Following this, the varying conditions of two opera productions led to two quite different recording approaches of immersive content for VR, a format that in other genres holds the common notion of standardised camera and microphone configurations. However, these two examples demonstrate the need to consider the individual requirements of the specific content as well as its associated genre. Whereas an opera recording in a legacy format (i.e. two-dimensional 16:9 aspect ratio) leaves it to the film director to edit camera cuts and foci of the action, VR hands over such responsibility to the spectator/user. It is thus even more important to suggest the user with specific contextual functions respective to the visual and acoustic domain of the presentation to establish guiding markers of the diegesis between the meta-virtual (i.e. the depiction of the stage performance) and its (infra-) virtual realm informed by the nature of the physical environment (i.e. the concert hall). If made aware of these roles on an implicit level during initial contact with the content (i.e. the overture of "Croak"), users are more likely to

recognise the meaning structures as well as the extend of possible actions on different layers of abstraction (i.e. what's to be depicted on stage versus what physical setting it is being staged in). Giving users of VR an initial understanding of the relevance and modal representation of each of these realms builds an authentic representation of a truly immersive experience that is more than a documentation of its original performance.

The volumetric simulations of "CroakVR" enable users to leave their audience seats and explore the hall while experiencing different close-up perspectives of the staging. More importantly, 6DoF and volumetric simulations motivate audiences to explore what lies in-between physical recording positions and their associated artefacts. We believe that an artistic rendition of this virtual space, brought about by design but also by artefacts introduced from the 6DoF conversion, highlights the intricacies of recording methods. As any kind of recording (including traditional formats) introduces a distortion of perspective compared to its natural acoustic counterpart, the contrasting between artistic (6DoF 3D environment) and perceptually naturalistic (3D 360° video) variants of CroakVR enables audiences to appreciate the impact of artefacts onto the overall experience. Thus, rather than concealing the presence of artefacts, they may be used to enhance the experience towards the kind of absorption that defines any ambitious staging: To expand spaces from the mimetic natural towards the diegetic artificial.

3.3 "Climb" (2017): an immersive "Concerto"

The third offering of our VR lineup tried to account for the shortcomings and improve the efficiency of the production pipeline in use during the previous two recordings. Furthermore, it considered the time limitations of the setup inherent to the location (Abbey Road Studios, London; Studio 2) and occasion of the concert (FAST research network at an open event as part of Abbey Road RED). In contrast to "She" and "Croak", the performance "Climb!" does not belong to a representative dramaturgic genre per se, but rather showcases an affinity towards programme music. Despite its focus on a narrative as part of the conceptual and compositional structure, the audience finds itself as part of a concert setting that deviates from the normal by means of an interaction between the player and the piano. Because "Climb!" does not rely on a unidirectional performance paradigm but affords a constant change of focus between two agents, the recording plan had to account for two protagonists rather than one; that is the pianist and the motorised piano itself. If not presented with the auto-performing piano en-face, the audience would have difficulty in entrusting the performance as

being authentic. This is true especially in the light of the electro-acoustic setup that was used to simulate the weather conditions (i.e. filter effects applied onto the audio stream obtained from microphones positioned inside the piano). Naturally, under those conditions the audience would be presented with sound originating from the acoustic instrument as much as with the processed signal output by speakers. Thus, it was important to support the users'/audiences' awareness of the two interacting agents, where the first would impersonate the human nature of the climber while the other laid out the circumstances and challenges of the environment where artificial, superhuman powers (i.e. weather or natural adversaries/challenges) outreach human limitations (note: ironically, it appears to be the most efficient path to mimic nature in music by opting for a rendering of artificial otherness that lies beyond the human grasp). Hence, the user would have to witness the action on the piano keyboard itself to understand the struggle between these two intersecting worlds. In order to account for these requirements, a first-person perspective was chosen where the main camera (Insta360 Pro in 3D at 6K resolution) was positioned right above the pianists' head. This perspective places the action of piano keys and hands of the pianist as well as the overall happenings inside the studio (i.e. technicians, audience) at the centre of attention while the head of the pianist would be positioned in the periphery of vision at about 150-degrees vertical.

Ultimately aiming to allow the user to walk around the piano in 180 degrees towards the right of the VR environment we set out to put in practice a more efficient method to capture the information required for the methods tested as part of the alpha "CROAK VR" experience. In total, four stereoscopic camera rigs were positioned in a half circle around the piano. Apart from the aforementioned first-person perspective, the second camera rig was placed towards the right of the performer at 45 degrees (2x Kandao QooCam in 3D at a combined 8K resolution; custom mount), the third rig at 90 degrees (2x Kandao QooCam in 3D at 8K) and the fourth rig at 180 degrees (2x Kodak SP360-4K in a stereoscopic configuration). In addition, a further camera (Insta360 Pro in 3D at 6K resolution) was positioned within the ranks of the audience's seating area.

Similar to "CROAK VR", the in-development "CLIMB! VR" experience will enable users to switch and seamlessly blend between these camera positions. However, the exclusive use of stereoscopic cameras facilitates the process of deriving depth information for the application with displacement maps (for a detailed overview see Gladstone, 2018a; and the Stereo2Depth python

script, Gladstone, 2018b). A bespoke procedure of spatial synchronization (i.e. pixel-by-pixel mapping of depth maps between camera positions) of the footage allows us to benefit from the relatively high number of stereoscopic 360 cameras surrounding the piano. Hence, it is possible to combine these perspectives and blend between depth maps and their associated displacement vertexes from each position at a relatively low count of artefacts. The blending between positions occurs at a relatively low threshold when the user moves outside the hotspot area of an active camera position to the degree that 6DoF related artefacts along the horizontal plane outweigh artefacts from a neighbouring camera position.

A similar approach is chosen for the implementation of audio within the VR environment, though, in addition to the streams obtained from one 3rd-order Ambisonics microphone (Zylia ZM-1), we also devised two additional 1st-order microphones (Soundfield SPS200) to account for the multitude of perspectives. However, during post-production it was decided to pursuit a similar approach as in CROAK by isolating sound sources from the 3rd-order Ambisonics recording and placing them as sound objects within the virtual environment. These sound objects act as dry perspectives whereas the 1st-order microphones provide a phase-aligned reverberant signal depending on the distance the user has moved away from the grand piano.

By making use of volumetric simulations, "Climb! VR" will allow its users to roam around the piano and experience different close-up perspectives of the performance, as in following the key action right above the shoulders of the performer or stepping over to observe the piano hammers and strings or to look at the pianist and audience from behind the grand piano. More importantly, 6DoF and volumetric simulations enable users to explore what lies in-between these camera positions and disclose a more immersive experience than what was originally bargained for with traditional stereoscopic recording equipment.

4. CONCLUSIONS

As showcased by the examples above, it is evident that AR and VR technologies can be used to distribute music performances to wider audiences by enabling concert experiences beyond the physical walls of a concert hall. Moreover, this kind of recording techniques can amend the life span of musical compositions that would otherwise mostly rely on the impact of their live performances attended by a limited number of listeners. Instead of just replicating the stage performance, a VR recording can enable such spectator perspectives

that would not be possible in a regular concert setup, as well as to underline the dramatic arch of the narrative, visuals, and stage presence: such a recording should be treated as an artwork in its own right.

XR technologies have also a vast potential of providing technical and expressive means for music creators. On top of allowing visual and sonic alternatives that would not otherwise be possible (for example, the immersive operatic forms by Michel van der Aa) they can be used, for example, to facilitate complex musical notation (Santini 2018) or as pedagogical tools for learning new music (see Pianolens 2016). We believe that XR will form an increasingly significant part of the future music producers´, composers´ and performers´ tool kit, thus enriching the overall concept of classical music culture.

5. REFERENCES

Audionamix (2019) *ADX Trax Pro 3*. Audio Spectral Editing Tool.
https://audionamix.com/technology/adx-trax-pro/ (retrieved 19 March 2019).

Gasselseder, H.-P. (2015) The Role of Agency in Ludoacoustic Immersion: Experiencing Recorded Sound and Music in Situational Context. *Audio Mostly*, Thessaloniki, Greece, 7–9 October 2015. ACM, New York.

Gladstone, J. (2018) *Stereo2Depth*.
http://pseudoscience.pictures/stereo2depth/pseudoscience_stereo2depth_1.8.5.zip (retrieved 19 March 2019).

Gladstone, J. (2018) Use depth maps to create 6DoF in Unity. *Immersive Shooter*.
https://www.immersiveshooter.com/2018/01/23/use-depth-maps-create-6dof-unity (retrieved 19 March 2019).

Insta360 (2017) *Insta360 Pro*. 360° 3D Camera for VR. https://www.insta360.com/product/insta360-pro (retrieved 19 March 2019).

Kallionpää, M., Greenhalgh, C., Hazzard, A., Weigl, D. M., Page, K. R. and Benford, S. (2017). Composing and Realising a Game-like Performance for Disklavier and Electronics. *NIME*, Copenhagen, Denmark, 15–18 May 2017.

Kallionpää, M., Chamberlain, A., and Gasselseder, H.P. (2018) Under Construction – Contemporary Opera in the Crossroads Between New Aesthetics, Techniques, and Technologies. *Audio Mostly*, Wrexham, UK, 12–14 September 2018. ACM, New York.

Kodak (2015) *Kodak SP360-4K*. 360° Camera. https://kodakpixpro.com/Europe/cameras/vrcamera/sp3604k (retrieved 19 March 2019).

Santini, G. (2018) Linear (Live Generated Interface and Notation Environment in Augmented Reality). *International Conference on Technologies for Music Notation and Representation* (Tenor 18), Montreal, 24–26 May 2018, 33–42.

Schütze, S. and Irwin-Schütze, S. (2018) *New Realities in Audio*. Boca Raton, FL: CRC Press.

Strasnick, E., Chambers, A., Jiang, L. and Tong, X. (2016) *Pianolens*: An Augmented Reality Interface for Piano Instruction. https://cs.stanford.edu/people/estrasni/otherproject s/pianolens.html (retrieved 19 March 2019).

van der Aa, M. (2018) "*Eight*" (description). https://www.vanderaa.net/eight (retrieved 19 March 2019).

Zylia (2019). *Zylia ZM-1*. Ambisonics Microphone Array. http://www.zylia.co/zylia-pro.html (retrieved 19 March 2019).