

Hyper360 Virtual Music Experience

Interactive Spatial Audio and AI Processing

Barnabas Takacs[†], Gergely Richter, Denes Kacsor, Zsuzsanna Vincze

PanoCAST/Drukka, Budapest, Hungary, btakacs@panocast.com

ABSTRACT

We present a virtual music experience of 12 analogue instruments as an interactive art piece. The audio from 22 original tracks was mixed into 7 switchable instrument groups, thus allowing viewers to experience true 3D spatial sound. By switching these groups ON and OFF during the performance they can educate themselves on how the song is built up while also playing along with a digital drum set placed in the center of the virtual space. Visuals include AI-processed 360° videos featuring the musicians surrounding the viewer. Specifically, they are digitally transformed as a tribute to Artificial Intelligence processing and showcase the outputs of cutting edge research on AI-based production techniques, such as real-time pose estimation, monocular depth estimation, and shot tracking extended to panoramic footage. Additional features include touchable particles as a gamification element (audio notes flying in space) which – when hit by the viewer – emit sounds timed precisely to make up the rhythm. Overall, the VRMusic system demonstrates several advanced computer vision and aural processing techniques wrapped into a single compelling environment. It is also a new type of interactive music experience and a novel genre of future VR content, where fans can play along with their favorite bands.

CCS CONCEPTS

• Computing methodologies~Virtual reality • Computing methodologies~Computer vision • Human-centered computing

KEYWORDS

Virtual Reality, Spatial Audio, Artificial Intelligence, Interactive Art

1. Description

- **Step#1** We recorded 12 instruments and 6 musicians arranged in a full circle with a 360° camera rig in the center resulting in **22 digital tracks** that were mixed in post-production to create 7 spatial audio groups for interaction.
- **Step#2:** In Hyper360 VRMusic experience these instrument groups arranged as 3D audio sources inside the environment allowing the user to interactively **turning them ON and OFF individually**, while listening to the performance. By rotating their head they can experience spatial audio like never before. In addition a **virtual digital drum set** was also added in the center, thus allowing the viewer to play the rhythm along with the band. Additional interactive elements involve **particle-based touchable audio notes that emit sounds** when hit and timed to reach the viewer precisely to make up the rhythm and delivering a complete audio-visual experience. This is demonstrated in *Figure 1*.
- **Step#3:** Next, the 360° footage was processed with state-of-the-art **AI-based computer vision** film production techniques extended to panoramic video. Specifically, first we applied **pose estimation** to detect people, then we stabilized the footage (that was necessary as camera was gently spinning) by **tracking performers** and thus creating an **automatic virtual camera path** [1] to re-render the video. To visualize the AI results, multiple videos were created overlaying Pose, 3D Face and Depth information, among others. The **panoramic depth** map was computed frame-

by-frame using a *monocular depth estimation* network [3] extended to handle equirectangular video [1]. Examples of multiple AI outputs are demonstrated in *Figure 2*.

- **Step#4:** The final experience includes a combination of these AI videos and renderings selectable by the viewer, while experimenting with 3D audio, playing along the music on the drum set, and plucking particle notes to match the rhythm. Screenshots are demonstrated in *Figure 3* and more information is available at this video: https://panocastdownloads.s3.eu-central-1.amazonaws.com/Hyper360VRMusic_SIGGRAPH2020_Demo.mp4

We argue that this interactive art installation showcases *a novel genre of future VR content*, where fans can play along with their favorite bands. The Hyper360 [2] solution acts as a generic framework for 360° video accompanied by true spatial audio. The complex analogue audio recording process used here, can be easily replaced with individual digital audio tracks from studio and concert recordings in order to quickly create similar experiences for any band.

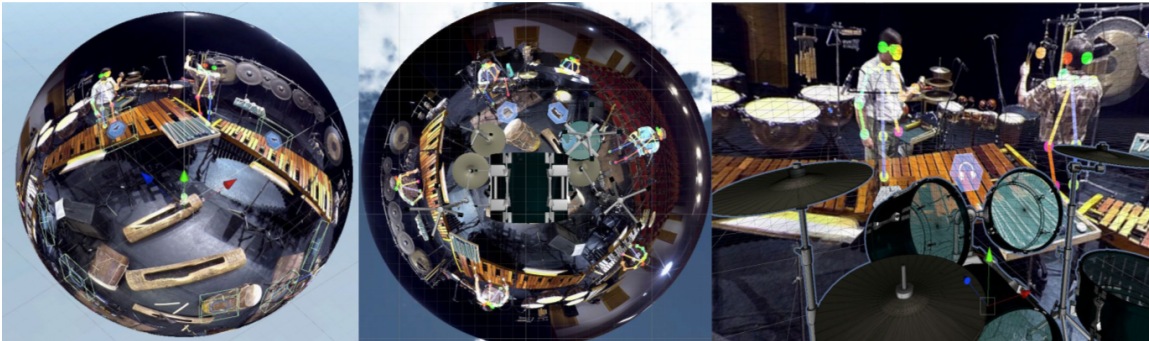


Figure 1: Spatial audio placement toggles in the 360° scene with a virtual drum set in the middle.



Figure 2: Examples of Artificial Intelligence processing (Left: pose tracking, Right: Spherical depth estimation)

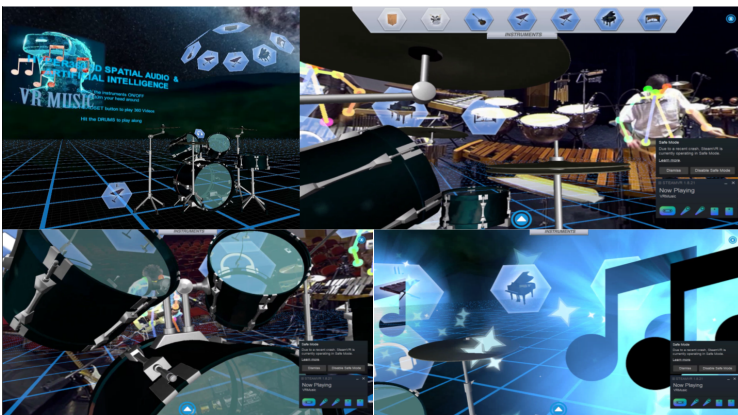


Figure 3: Example screenshots from the application.

ACKNOWLEDGMENTS

This work has received funding from the European Union’s Horizon 2020 research and innovation programme, grant n° 761934, Hyper360 (“Enriching 360 media with 3D storytelling and personalisation elements”).

REFERENCES

- [1] Barnabas Takacs, et. al. 2019. Hyper 360 – Towards a Unified Tool Set Supporting Next Generation VR Film and TV Productions in J. Software Engineering and Applications, 12, 127-148 (May 2019), 10 pages, <https://doi.org/10.4236/jsea.2019.125009>
- [2] <http://www.hyper360.eu/>
- [3] <https://ai.googleblog.com/2019/05/moving-camera-moving-people-deep.html>