

360° Video Database with Higher-Order Ambisonics Spatial Audio

Thomas Robotham², Ashutosh Singla¹, Olli Rummukainen², Emanuël Habets², Alexander Raake¹,

¹Audiovisual Technology Group (AVT), TU Ilmenau

²International Audio Laboratories Erlangen (A joint institution of the Friedrich-Alexander-University Erlangen-Nürnberg (FAU) and Fraunhofer Institute for Integrated Circuits (IIS))

DFG Deutsche
Forschungsgemeinschaft
German Research Foundation

Motivation

- New video coding algorithms (e.g. AV1 or Versatile Video Coding (VVC)) suitable for high resolution have been proposed
 - Challenge → There are not so many high quality publicly available datasets
- Most of SoA datasets are available without audio and are of short duration (10 s)
 - 10 s standard-set videos are too short to meaningfully investigate effect of simulator sickness and presence
- YouTube, Arte → Not available in high-quality versions
- Lack of datasets containing higher-order Ambisonics audio and head-tracked audio rendering
- Lack of availability of audiovisual 360° databases

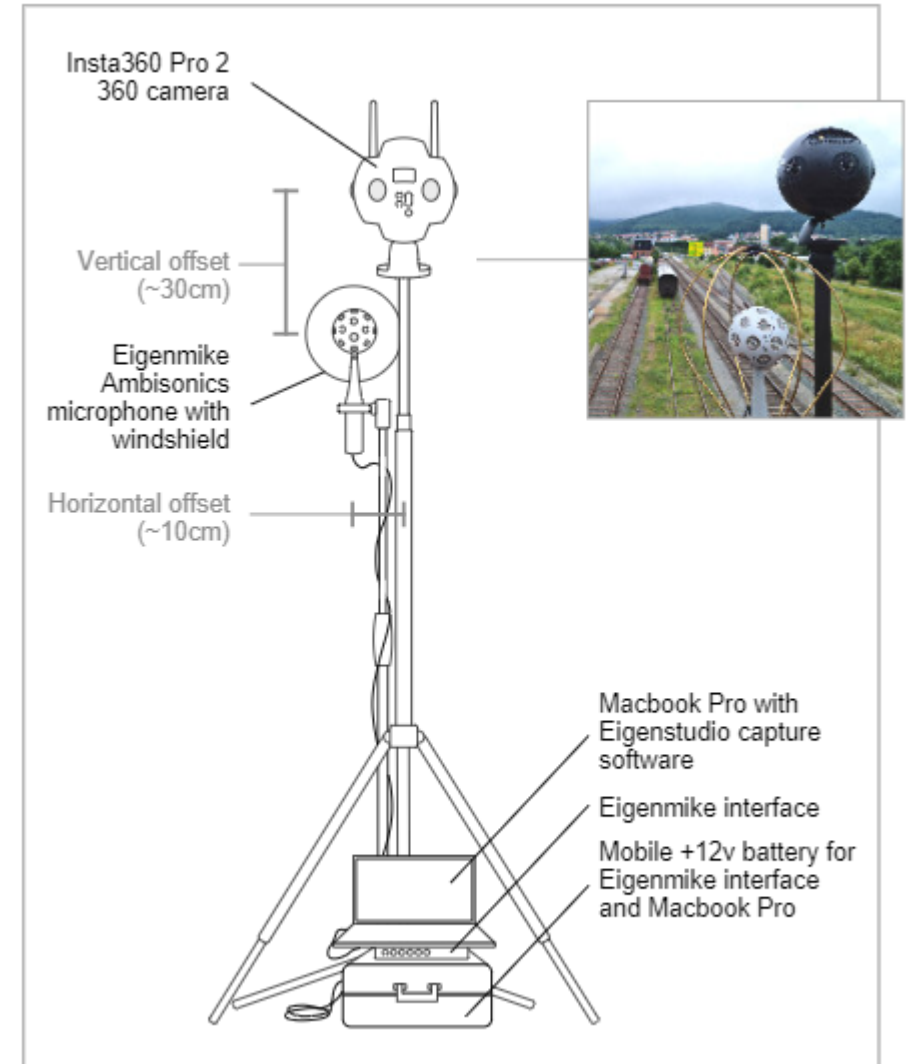


Existing 360° Video Datasets

Author	Resolution	Duration of a Video	Audio	Video Count	Source
Abbas et al. [1]	8192×4096	10 s	No	6	GoPro Omni
Asbun et al. [2]	8192×4096	10 s	No	4	GoPro Omni
Duan et al. [8]	4096×2048	15 s	Yes	10	Insta360
Li et al. [12]	No Info	29-668 s	Yes	73	YouTube, Vvideo, Facebook
Liu et al. [14]	8192×4096	30 s	No Info	12	AURA 360
Mahmoudpour et al. [16]	3840×1920	6-12 s	No Info	15	GoPro Fusion & YouTube
Schwarz et al. [24]	3960×1980	10 s	No	1	Nokia OZO
Sun et al. [26]	8192×4096 & 4096×2048	10 s	No	9	Sony A7RII & GoPro
Xu et al. [27]	2880×1440 & 7680×3840	20-60 s	No	48	YouTube & VRCun
Zhang et al. [28]	4096×2048	10 s	No Info	16	No Info
Maugey et al. [17]	3840×1920	60-240 s	No Info	8	Samsung Gear 360
Nasrabadi et al. [21]	3840×1920	60 s	No	30	Samsung Gear360 & YouTube
Mazzola et al. [18]	3840×2160	15-90 s	No Info	16	Garmin VIRB 360
Chao et al. [5]	3840×1920	25 s	First-order Ambisonics	15	YouTube

QoEvaVE 360° Video Dataset

- 11 Videos
 - Shot in Ilmenau, Nürnberg
 - Capturing real-life scenes
 - Audio and video recorded separately
- **Video:** Insta360 Pro2
 - Static camera
 - Resolution: 7680 × 3840
 - Duration: 60 s → 10 s, or 20 s or 30 s
 - Framerate: 60 fps
- **Audio:**
 - Eigenmike with windshield
 - 4th order Ambisonics audio



Camera and Microphone Configuration

Source Contents – Overview



Some Sample Scenes



Some Sample Scenes

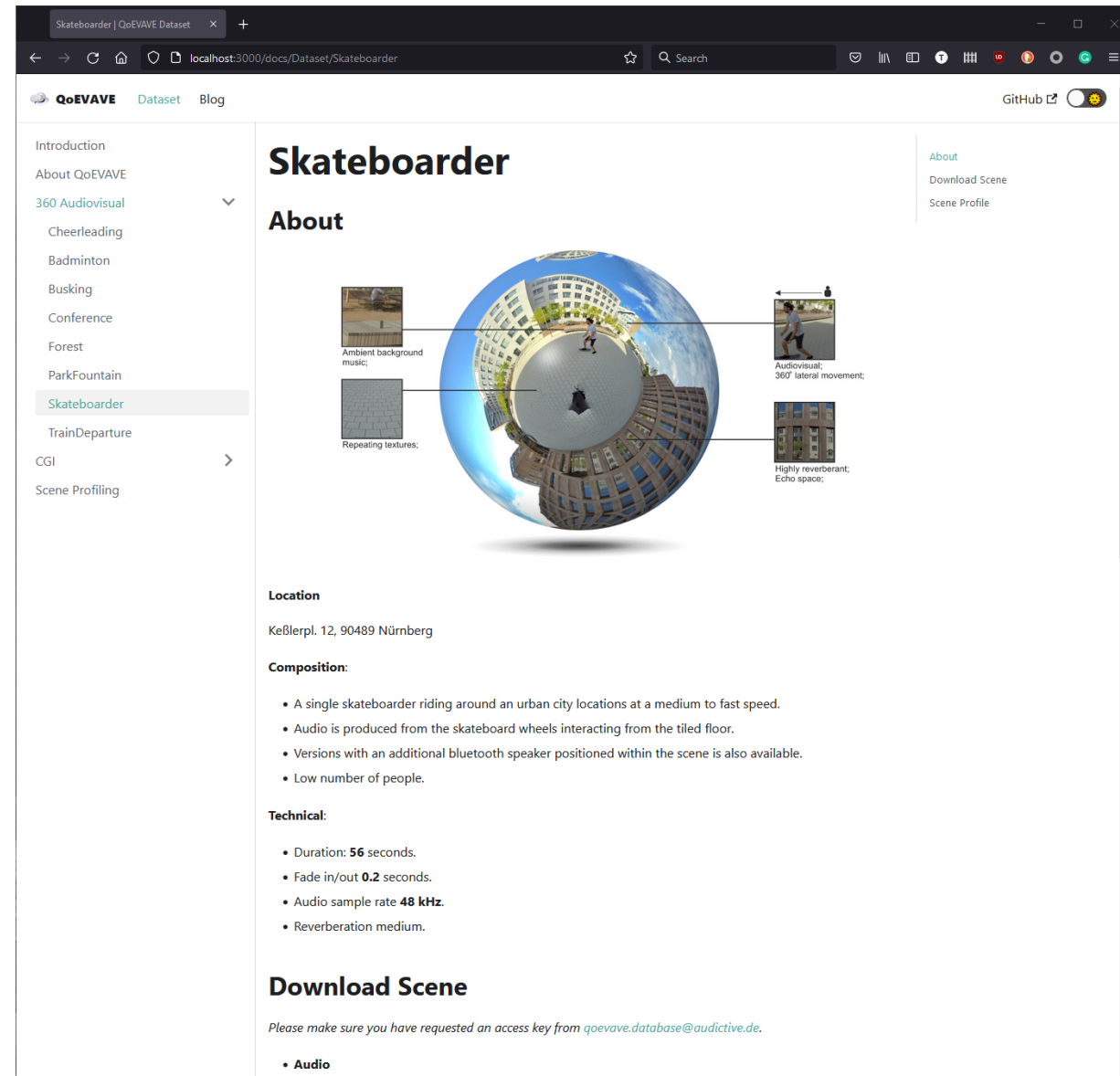


Some Sample Scenes



Dataset Hosting

- Dataset (Website) → will be made publicly available
 - Various framerates, resolutions, or Ambisonic orders
 - Audio and visual content is available independently
 - Also, single file in mp4 container
- YouTube
 - Audio: First-order Ambisonic
- These scenes are first iteration of ongoing production cycle
- Further extended and augmented by VR scenes
 - 6 DoF



The screenshot shows a web browser displaying the 'Skateboarder' dataset page on the QoEVAVE website. The page layout includes a navigation menu on the left with categories like 'Introduction', 'About QoEVAVE', '360 Audiovisual', and 'Skateboarder' (which is highlighted). The main content area features a large 360-degree panoramic image of a skateboarder in an urban environment. Surrounding this central image are four smaller thumbnail images with labels: 'Ambient background music;', 'Repeating textures;', 'Audiovisual; 360° lateral movement;', and 'Highly reverberant; Echo space;'. Below the main image, there are sections for 'Location' (Keßlerpl. 12, 90489 Nürnberg), 'Composition' (listing details about the skateboarder, audio production, and speaker), and 'Technical' (listing duration, fade in/out, audio sample rate, and reverberation). A 'Download Scene' section is also present, with a note about requesting an access key from qoevave.database@audictive.de.

Thank you for your attention!
We look forward to your questions.



Prof. Dr.-Ing.
Alexander Raake



Dr. Olli Rummukainen



Prof. Dr. ir.
Emanuël Habets



Thomas Robotham, M.Sc



Ashutosh Singla, M.Sc