Visualizing and Sonifying Scenes of Potential Future Urbanism in Shenzhen's Sea World in "High Sections / Low Leaps"

Marcel Zaes Sagesser[*†]
marcel.sagesser@fhstp.ac.at
St. Pölten University of Applied Sciences
St. Pölten, Austria

Tristan Braud[*]
braudt@ust.hk
The Hong Kong University of Science and Technology
Hong Kong, China

Gyuwon Sylvia Lee[*]
gwsylvia.lee@gmail.com
Freelance Artist
Seoul, South Korea

Zhen Wu[*]
zwuch@connect.ust.hk
The Hong Kong University of Science and Technology
Hong Kong, China

Figure 1: Screenshot of the artwork, cross media composition in Unity, visualizing technological urbanism, by the authors, 2025.

Abstract

In Shenzhen's Sea World in China's Greater Bay Area, on reclaimed land, ocean-side promenades are set in front of impressive glass towers. Next to it lie container port, construction sites, and densely populated older areas. A rapidly built vision of Shenzhen's future co-exists with traces from its various pasts. "High Sections / Low Leaps" is an audiovisual artwork that reconstructs actual and imagined development phases of China's Greater Bay Area, one of the fastest developing areas on the planet. This artwork element, with the subtitle "Shenzhen's Sea World Ocean Mall," is an interactive scene rendered in a game engine, presented to the audience on two screens, showing otherworldly office towers that grow and shrink. 3D modeling, AI-generated facade elements, and spatial audio are employed to create dynamic audiovisual cityscape speculations. A camera tracks the number of spectators in proximity and switches the scene from past to future phases, allowing for subtle control over the work's temporality. The artwork investigates how rapid urban development, such as in the Greater Bay Area, can be rendered perceptible for a public, fostering reflection on the relationship between new technologies (including GenAI) and imagined (and built) urban futures. At the intersection of computer graphics, data visualization, urbanism, creative technologies, spatial audio, and generative AI, this approach argues for a reflection on the consequences of human-technological collaboration in view of desired humane rather than technological outcomes. The pilot city Shenzhen holds significance for cities globally.

CCS Concepts

• **Social and professional topics** → Cultural characteristics; *Geographic characteristics*; • **Human-centered computing** → **Visualization**; **Geographic visualization**; • **Applied computing** → **Media arts**; **Sound and music computing**; **Fine arts**.

Keywords

Urbanism, Speculative Futures, Creative Technologies, GenAI, Interactive Arts, Spatial Audio

ACM Reference Format:
Marcel Zaes Sagesser, Tristan Braud, Gyuwon Sylvia Lee, and Zhen Wu. 2025. Visualizing and Sonifying Scenes of Potential Future Urbanism in Shenzhen's Sea World in "High Sections / Low Leaps". In *Proceedings of the 18th International Symposium on Visual Information Communication and Interaction (VINCI 2025), December 01–03, 2025, Linz, Austria.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3769534.3769591

[*]All four authors contributed equally to this research.
[†]Corresponding author

VINCI 2025, Linz, Austria

ACM ISBN 979-8-4007-1845-8/25/12
https://doi.org/10.1145/3769534.3769591

1 Artist Statement & Concept

How are urban development, new technologies (including AI), and imaginative futures linked? How could more humane urban environments be achieved for our future cities? The project addresses these research questions through a creative-speculative, interpretative reconstruction of Shenzhen's Sea World district in a virtual environment. It visualizes the consequences of letting technologies, rather than humans, design our urban environments. Letting spectators interact with different development phases, presented as otherworldly visualizations of imagined pasts and futures, renders public awareness of the theme. The work's goal is to foster dialogue about desired urban futures and the complex, often under-reflected role of our contemporary technologies, including generative AI. For thematical reference, see also [1, 2].

2 Creative Process

We prompted AI to give us visual renditions of potential futures. The image creation process mainly utilizing Midjourney involved a diverse selection of keywords such as "oceanfront," "high-tech parks,", "development", "port city" as well as original photos taken from the site as image prompts. For animated versions, Genmo's new model Mochi 1 and Sora were used with different approaches. On Genmo, we emphasized dynamic camera movement through the text prompt in addition to the prompts for images. With Sora, we focused on keywords such as "process" and "transform" in addition to the original photos of the location as image prompts in the storyboarding feature to visualize change of location over a timespan. Overall, Genmo showed the odd morphing effect for animation we were aiming at but at a loss of details and clarity while Sora generated videos that were more realistic but less surreal and diversified. With this method, 30+ short animations created.

Figure 2: Installation with many spectators, triggering the scene to show land reclamation with rising buildings.

In the next step, we rebuilt the present Sea World district in the game engine (Unity) and added speculative 3D models into the scene for its past and future. The scene is divided in four phases, each corresponding to a district's development phase. The AI-generated video clips were added as texture over each building. All videos are 7 seconds long and loop with synchronized playback with the game time. To infuse more life and emphasize the district's ever-evolving nature, we vary the height of the building over time, synchronized with the video playback. A day-night cycles establishes a larger rhythm of change (see figure 1).

3 User Interaction: Screen & Sound

The two channels of video, showing the same neighborhood and the same day-night cycle from different camera positions, are installed in the gallery (figures 2 and 3). The artwork presentation involves two monitors, a computer, two loudspeakers and a camera. The monitors are aligned with the digital view (Unity camera positions) they are displaying as "windows to another world." The camera captures the rough number of users in proximity to the work. In Unity, we interpolate between different phases or urban development depending on how many users are close by. The two screens display the same moment of the urban space from different perspectives, providing the audience with a detailed perception of the urban changes. The sound, provided to the audience via loudspeakers, consists of spatial audio rendered in Unity, matching one of the two camera perspectives displayed. Field recordings from Sea World are preprocessed with filtering and granular synthesis to achieve a techno-organic blend. The soundtrack is made intractable with filters that change as more or less audience members are close by. The sound equally presents speculative soundscapes that mix real with imaginaries.

Figure 3: Technical setup of the installation with LCD displays, loudspeakers, computer, and camera.

4 Privacy concerns

Our camera captures will be processed through a local computer vision model, utilizing YOLO or a locally trained recognition system to estimate the total number of people nearby. The data will not be sent to a cloud service, nor be stored on the device. A printed notice informs visitors about the camera detection and the data handling.

Acknowledgments

This work has been partially supported by the Southern University of Science and Technology, School of Design, Shenzhen, and partially by SeeD Creative Commune in Shenzhen-Longgang.

References

[1] Federico Caprotti and Dong Liu. 2020. Emerging platform urbanism in China: Reconfigurations of data, citizenship and materialities. *Technological Forecasting and Social Change* 151 (Feb. 2020). https://doi.org/10.1016/j.techfore.2019.06.016

[2] Dongjie Wang, Yanjie Fu, Kunpeng Liu, Fanglan Chen, Pengyang Wang, and Chang-Tien Lu. 2023. Automated Urban Planning for Reimagining City Configuration via Adversarial Learning: Quantification, Generation, and Evaluation. *ACM Trans. Spatial Algorithms Syst.* 9, 1 (2023). https://doi.org/10.1145/3524302