Towards an Immersive Audio-visual Model of Copacabana for Future Urban Studies

Jonas HECK[1], Christian DREIER[1]

[1]Institute for Hearing Technology and Acoustics, RWTH Aachen University, Kopernikusstr. 5, 52074 Aachen, Germany

jonas.heck@akustik.rwth-aachen.de

Abstract. *Urban design strives for a positive impact on people's well-being. To study the person-oriented perception of future urban design concepts, we propose to use immersive audio-visual representations of the virtual space. This work presents the procedure to create a photogrammetry-based model of the iconic scenery of Copacabana, Rio de Janeiro, Brazil. The proposed procedure delivers a description of the photographic data acquisition. The following post-processing including alignment, filtering, and noise removal results in a textured mesh. The acoustic characteristics of the urban environment are depicted. Finally, a simplified model for acoustic simulation is derived. Both the visual and the acoustic model together can be used as a framework for multi-modal perceptual studies in complex urban environments under controlled laboratory conditions.*

Keywords

Auralization, digital twin, audio-visual model, urban soundscape, urban planning.

1. Introduction

The interest in health-promoting and person-oriented urban environments steadily increases. A way to approach this development is to study people's perceptual assessment of design proposals and to involve people into the planning process of those. As acoustics plays an important role in this context, there is great potential for the use of auralization in virtual reality. The concept of auralization allows to separately control sound sources, sound propagation and receiver configurations [1]. Complemented with an immersive three-dimensional visual representation, those models can serve as a plausible environment of the urban space to study people's perception and preferences under laboratory conditions.

Visual models of concert halls and lecture rooms have been derived from photogrammetry earlier [2]. Furthermore, the procedure has already been applied to an outdoor scenario, i.e., a green space in Aachen, Germany [3, 4]. While in the case of the green space, the area has a size of about 170*110 m and is usually visited by few pedestrians only, the present paper aims to depict the procedure in a larger and more complex environment with many moving sources such as pedestrians, bicycles, cars, buses, airplanes and ships.

For the photogrammetry, first, a considerable amount of photographic data needed to be acquired. Therefore, several photography walks were performed. The photography walks have been complemented with 360° video and simultaneous audio recordings in first-order Ambisonics format. Two models, for both the visual and the acoustical representation of the space, have been derived. The derived models as well as the corresponding audio recordings of important sound sources and the overall soundscape serve for documentation purposes. Furthermore, they can be used for future studies of urban soundscapes and urban design concepts.

2. Photogrammetry

For Copacabana, no big amount of sufficiently high-quality photographic data were available that could be used to build a photogrammetric model of high resolution. Therefore, own photographs were taken between November 29 and December 10, 2021. The path of the photography walks is mapped in figure 1. The camera was a Canon EOS R5 body with a Canon RF 24-105/4.0-7-1 IS STM lens. The photographs were taken with a resolution of 8192*5464 px, either at head-height (hand-held) or on a tripod, as depicted in the figure 1.

Fig. 1. Camera positions. Blue: Hand-held at head height. Red: On a tripod at 8–12 m height, rotating 360 degrees. Cyan (dotted line): Both hand-held and tripod.

The photographs were post-processed with the software Agisoft Metashape (version 1.8.3). It aligns multiple pictures by triangulation and allows to generate point clouds, meshes as well as textured object models [5, 6]. After manually removing some ineligible photographs, in detail, the post-processing contained the following steps, cf. figure 2.

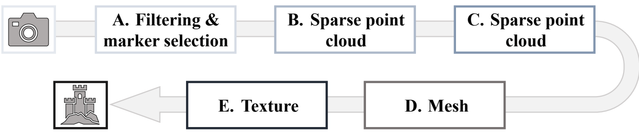

Fig. 2. Photogrammetry workflow.

A. Filtering and marker selection:
   - Excluding pictures with a non-sufficient quality.
   - Markers were manually positioned on, e.g., manhole covers and road signs to align those pictures that were not automatically aligned in first place.
   - Masking was applied to remove parts of pictures compromising the point cloud quality (mainly sky).
   - Further filters, i.e., ”reconstruction uncertainty”, ”projection precision”, ”reproduction error”, and optimization of the camera positions.

B. Sparse point cloud

C. Dense point cloud:
   - Bounding box to exclude distant outliers.
   - Noise removal according to the confidence level.
   - Manual removal of ghost objects (cars, trees, road signs).
   - Minimum spacing between points to reduce overdetermined areas.

D. Mesh generation from dense point cloud:
   - Interpolation to close remaining holes.
   - Smoothing to remove too complex polygon structures.

E. Texturing of the mesh with a texture map of 4096x1 px.

The quality of the reproduced model was so far insufficient for an visual immersion due to areas of no or too sparse picture information. Furthermore, the holes in the mesh prevent the use of geometrical acoustic simulation techniques. To close the model, we combined the model derived from the presented terrestrial photogrammetry with a photogrammetry-based model derived from Google Earth[1]

[1] Accessed on 2022-06-13.

Fig. 3. Combined photogrammetry model (terrestrial + flyover photographs) with texture.

flyover captures of the waterfront. The result of this amendment, which is only incorporated into the model close to the waterfront so far, is depicted in figure 3.

3. Acoustic Topography

The choice of the investigated space was made with the help of our project partners from the Federal University of Rio de Janeiro (UFRJ) who could support us with local knowledge of typical traffic flow etc. This helped to prepare and schedule not only the audio recordings but also the photography walks, especially with regards to the natural lighting conditions. An overview of the acoustic topography and characteristic acoustical features of the wavefront of Copacabana is shown in figure 4. The detailed description of the urban sound environment as well as its characteristics follows below.

The space was selected for its special topography which results in key features for acoustic simulations. The main factors influencing acoustic sound propagation are (see also figure 4):

I. Facades of the 6-8 story buildings along Avenida Atlântica with various degree of absorption and scattering.

II. Pavement and asphalt of Avenida Atlântica, its promenade + bicycle lane between street and beach, as well as the opposite sideway between street and buildings.
   - Some smaller objects that influence sound propagation, i.e., kiosks with sunshades, chairs and tables.
   - Parking, stopping and driving cars, buses and trucks.

III. Highly absorptive sand at the beach.

IV. Wavefront, idealized as an acoustic line source.

V. Highly reflective water surface on the sea.

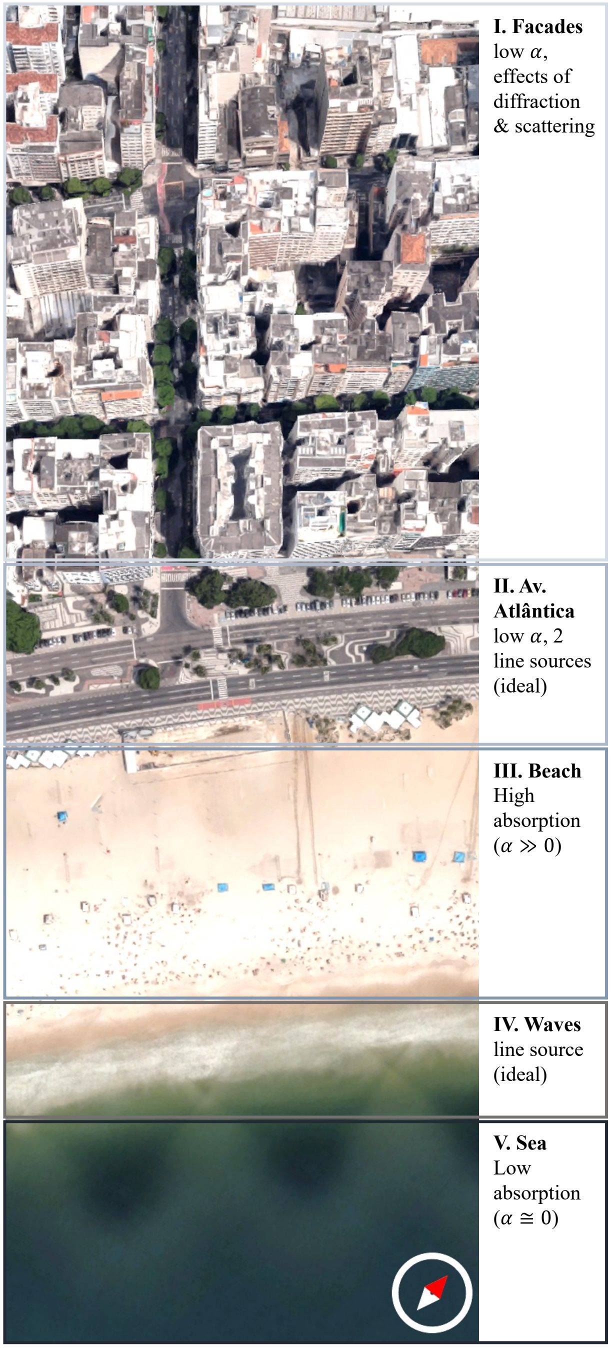

Fig. 4. Aerial view (Source: https://www.google.com/maps, accessed on 2022-07-26) of the city block under study, depicting its acoustical features.

The following sound sources could be observed at Copacabana:

- Road traffic noise (cars, buses, trucks, motorcycles, alarm signals of emergency services like ambulance and police).
  - Heavy motorized traffic on Avenida Atlântica on both lanes, in the morning mainly in northern direction.
  - One-way road Avenida Nossa Senhora de Copacabana: Faster traffic with lots of buses passing.
- Air traffic (small airplanes and helicopters).
- Music from the bars/kiosks along the beach.
- Generators at the beach which are used to pump the water to provisional showers.
- Ocean sounds (waves and sea spray as well as birds).
- Babble of voices and voices of vendors.

While it is, in principle, possible to automatically reduce the mesh size (the number of polygons that form the mesh), we could not successfully keep important geometries stable and exact. Mainly, the details of facades, such as ventilators, windows, balconies, balustrade and projections, were not maintained. Furthermore, the mesh geometry is too complex for efficient acoustic simulation. This is why a new CAD model for the use in real-time auralizations was sketched that complies with the requirement of a reduced number of mesh polygons. It was manually derived from the photogrammetry mesh. The derived model uses dimensions measured from the photogrammetry model (cf. figure 3) and was built with simple rectangular blocks. The CAD sketch of the model can be seen in figure 5.

The materials were estimated from the photographs and the corresponding acoustic properties were defined on the object surfaces. From this on, the acoustic model is ready to be used with established simulation tools, e.g., to auralize synthesized pass-by noise in urban environment [7].

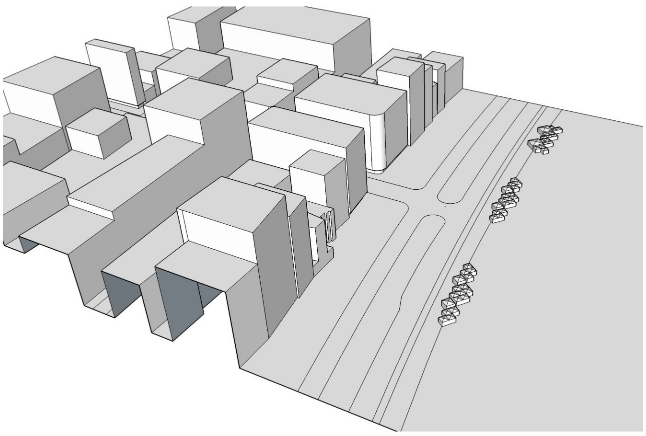

Fig. 5. CAD model of the urban environment at Copacabana of reduced complexity for acoustic simulation.

4. Conclusion and Outlook

Firstly, a procedure of capturing and post-processing a complex urban environment for photogrammetry modelling was illustrated. The photogrammetry model can serve as a virtual environment for visual immersion. The drawbacks we encountered during the post-processing should be briefly discussed:

- For the camera configuration on the high tripod, we slightly tilted the camera to not only capture the horizon but mainly the scene of interest. This could sometimes have resulted in alignment problems because the horizon may not always have been clearly recognizable as a landmark.
- Through adding flyover pictures to our model, it was possible to close the previously perforated mesh. But the resolution of geometrical details, especially of the facades, as well as the texture is much lower than from the high-resolution pictures taken by the authors.

Future work will contain a refinement of the post-processing steps with more detailed picture basis (higher amount of data points), other options for texturing the mesh, the incorporation of flyover photographs at multiple stages and the testing of filtering algorithms to delete disturbing objects from the photogrammetry model.

Secondly, the acoustic topography of Copacabana was described. With this and the visual model, an acoustic model for geometrical acoustic simulations was manually derived. In future, we will have to work on finding mesh reduction algorithms which maintain important geometric features of building facades that are of importance for the acoustic simulation while reducing the number of polygons significantly. This would allow us to reduce the manual effort to derive the acoustic model from the photogrammetry model.

Both, the visual and the acoustic model can be used together in interactive audio-visual experiments. Those experiments allow us to study individual noise perception or the influence of architectural design on humans on urban scale. The final models can be scaled down or complemented in various degrees of detail accounting for plausible immersion and for the respective purpose.

Acknowledgements

The topic has been supervised by Prof. Michael Vorländer, Chair of Technical Acoustics at RWTH Aachen University, and supported by the the German Academic Exchange Service (DAAD). The authors acknowledge the kind help from the exchange counterparts from the Federal University of Rio de Janeiro (UFRJ), especially the on-site support from Prof. Mariane Rembold Petraglia and Prof. Julio Cesar Boscher Torres. Furthermore, the authors like to thank Mr. Artur Fuchs and Dr. Josep Llorca-Bofí for their help during the photogrammetric post-processing.

References

[1] VORLÄNDER, M. *Auralization. Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. 2nd edition. Springer Cham, 2020. https://doi.org/10.1007/978-3-030-51202-6.

[2] LLORCA-BOFÍ, J.; WITEW, I.; REDONDO, E.; VORLÄNDER, M. 3D modelling photogrammetry to support acoustic measurements and derive geometries for simulation. In: *Proceedings of the Institute of Acoustics (Auditorium Acoustics 2018)*, 2018, 40 661-666.

[3] LLORCA-BOFÍ, J.; VORLÄNDER M. Multi-Detailed 3D Architectural Framework for Sound Perception Research in Virtual Reality. *Front. Built Environ*, 2021, 7:687237. https://doi.org/10.3389/fbuil.2021.687237.

[4] LLORCA-BOFÍ, J.; DREIER, C.; HECK, J.; VORLÄNDER, M. Urban Sound Auralization and Visualization Framework — Case Study at IHTApark. *Sustainability*, 2022, 14(4), 2026. https://doi.org/10.3390/su14042026.

[5] MIKHAIL, E. M.; BETHEL, J. S.; MCGLONE, J. C. *Introduction to Modern Photogrammetry*. John Wiley & Sons, 2001. ISBN: 978-0-471-30924-6.

[6] AGISOFT LLC. *Agisoft Metashape User Manual: Professional Edition. Version 1.8.*, St. Petersburg, Russia, 2022. URL: https://www.agisoft.com/downloads/user-manuals/.

[7] DREIER, C.; HAHN, J.; HECK, J.; LLORCA-BOFÍ, J.; VORLÄNDER, M. Real-time vehicle pass-by noise synthesis for urban sound auralization. In: *Proceedings of the 24th International Congress on Acoustics*, Gyeongju, South Korea, 2022.