PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Progressive region-of-interest filtering for urban sound auralization applications with multiple reflected and diffracted propagation paths

Jonas STIENEN; Michael VORLÄNDER

Institute of Technical Acoustics, RWTH Aachen University, Germany

ABSTRACT

When auralizing urban environments, two major challenges can be identified for the propagation simulation: the extent of the built environment and the wave-based propagation features, above all diffraction into the shadow region. It is in the nature of open space that sound can propagate long distances. For the auralization of city acoustics without a-priori knowledge on the sound source characteristics, huge data must be taken into account. With the increase of freely accessible open format databases providing layers with simplified geometric models of entire landscapes including buildings, a solution to create virtual acoustic environments based on this data is desirable. In this attempt, a progressive broad-band propagation simulation model based on Geometrical Acoustics is proposed that primarily exploits the frequency-independent energy loss subject to geometrical spreading of spherical waves for the spatial filtering of the geometry input. An ellipsoidal hull representing a region of interest is constructed from the source and receiver position as well as sound power and sensitivity, respectively. Subsequently a fast filtering algorithm with linear complexity removes approximately irrelevant geometry for each propagation step releasing the propagation algorithm from executing ray-polygon intersection tests on the entire geometry dataset and speeding up the propagation simulation significantly.

Keywords: Virtual Acoustics, Outdoor Sound Propagation, Geometrical Acoustics

1. INTRODUCTION

The auralization of urban environments is an aspired aspect for the forecast of our future city soundscape. The increasing generation and integration of digital geometric data as well as the progress to make federal procedures transparent lead to freely available online data, like the street and building maps published and maintained by the cadastral office of North-Rhine Westphalia, Germany[1]. Semantic data layers of various context, which enrich plain mesh files with surface material, traffic statistics, visitor counts, etc. form a vital information basis that invite multidisciplinary research parties to simulate current and future aspects of city life.

Obviously, the feasibility for the prediction of dynamics of sound is one of the questions raised. For environmental noise assessment, it has been demonstrated that semantic geometrical data can be integrated into acoustic prediction models, for example by Lu et al. (1) and Kumar et al. (2). Those modelling approaches integrate noise immission over a long-time span – usually hours – and present results in perceptually weighted energetic frequency bands. Sound sources are commonly represented by point sources and line sources producing a steady-state sound power spectrum. Judgements of the compliance with legal directives of these well-established noise levels can only be justified by consulting experts.

Auralization, in contrast, addresses the auditory perception and is intuitively comprehensible by layman, hence may be an effective tool for public information campaigns. In this sense, the environment is not evaluated to generate exhaustive noise maps, but focus must be laid on sensitive receiver points, ideally arbitrarily chosen by the user in three-dimensional space. Relevant objects in the vicinity of the sound source and the receiver must be considered, hence especially the geometry of the neighborhood region is of interest. A major difference between noise assessment and auralization is the fact that the latter does not average acoustic values over time but produces time series of pressure, an audio stream, that can be used for acoustic presentation of the scenario. Hereby, the time-variant acoustic transmission of each sound source is simulated by superimposing propagation paths to the

[1] Geoportal.NRW, https://www.geoportal.nrw, accessed Mai 2019

receiver via the built environment. Dynamics of the scenario are adapted and are processed by means of digital signal processing (3). In theory, effects like distant highway noise immission appearing as a steady-state noise spectrum following the spreading loss of a line source is inherently produced by physically modelling all single sound events. Ground effects and level corrections must not be considered, as the processing of coherent sound from a sound source is applied. However, in a dynamic acoustic environment the simulation of propagation paths and the representation of such paths in the DSP network is demanding and requires the capability to consider higher order reflections and diffractions in order to maintain an output that is free of perceivable artifacts from the time-variant adaption (4).

2. AURALIZATION IN THE CONTEXT OF ENVIRONMENTAL NOISE

Auralization is the process of making sound audible by physical means and with a resolution matching or exceeding the auditory processing capability. A distinct challenge is the high-quality requirements laid upon the input data. A virtual acoustic scenario is recreating a situation where a single user perceives sound propagated through the environment that is originating from one or many sound sources. In the context of environmental noise, this requires a *precise spatial mesh* of buildings as the user can approach facades closely and move around corners where effects like sound shadowing is prominent and even small movements can significantly alter the result. In general, the possibility to represent buildings is of high resolution and surpasses the highest wavelength of interest by far, i.e. structural details below $\lambda = 7mm$ can easily be represented, as also Kamrath et al. pointed out fittingly (5). The available data from online databases can readily be included without further refinement for auralization. Even level-of-detail concepts that only rudimentarily approximate buildings (e.g. by coarse bounding hulls) are practical in acoustics, if they do not significantly diverge from a visual model in an audio-visual Virtual Reality application.

In contrast to the spatial resolution, the *temporal resolution* of current noise assessment measures is relatively low, as it integrates over time in a magnitude of hours. Input data from sound sources and results from noise maps and suchlike will hardly be reasonable to integrate in auralization applications, mainly because the energetic values have averaged out all temporal details. The result of an auralization application delivers a pressure signal over time at a receiver location including all temporal variations a human being is sensitive to. However, individual results implementing outdoor sound propagation for environmental noise are of use. The individual character of a perceived sound, i.e. the emitted signal, is carrying information that is associated with naturalism, realism and authenticity (6). Modification of this signal by filters approximating the propagation effects, even without phase information and with low spectral resolution, maintains the typical sound and achieve successful auralization results (3). Increasing spaciousness and presence is then achieved by appropriately adding the environmental influence, which is the prime target of auralization. Apart from the high demand on the input signal of sound sources, the determination of the transmission within an urban scenario must be computed correctly in the sense that the filtering of the signal emitted by the source reflects the expected effects caused by the surrounding, like reflections and diffractions. Additionally, relative movements must be accounted for, namely the Doppler shift either caused by source movement or by receiver movement, at least when relative distance changes result in a perceivable frequency shift (7). In the context of environmental noise, implementing this dynamic component is one of the most challenging aspects and requires time-variant processing for each individual path.

3. PROGRESSIVE REGION-OF-INTEREST FILTERING

3.1 Breaking down the number of propagation paths

One of the greatest challenges of sound propagation simulation is the sheer size of the input data representing the built environment of urban areas. Resources like the OpenGeoData NRW[2] provide files including areal patches of 1000 m by 1000 m extend, each including several thousand buildings in case of densely populated regions. On the one hand, auralization demands arbitrary positioning of acoustic entities, on the other hand it is apparent that only very few structures in the vicinity of these objects actually play a role for the propagation of sound waves. Hence, a practical procedure to find

[2] https://www.opengeodata.nrw.de, accessed May 2019

and separate these structures is obviously of major interest. This requirement becomes even more important for acoustic algorithms that tackle early reflection simulation problems because they usually scale exponentially with reflection/diffraction order to the base of polygon count. Therefore, a method to filter a region of interest from the whole database is desirable, that automatically includes all relevant objects (polygons) and efficiently dismisses irrelevant data. If the decision process is lightweight yet driven by acoustic semantics and sorts out objects as soon as possible, the potentially more sophisticated subsequent simulations gain a large profit from adjusted input data without irrelevant areas. 'Irrelevant' in this context means, that the large part of the acoustically perceivable paths should still be detectable within the subset, while those being most likely undetectable by the human auditory processing should be culled. Such a perception-progressive filtering trait can't be avoided since theoretically the dynamic range of acoustic transmission is unlimited, for example, endless paths can be found that bounce between parallel facades or circle around buildings. However, it is clear that on the emission end, the sound power level of common sound sources appearing in urban environments have an upper limit, and on the receiving end, a certain detection threshold can be assumed. For humans this limit may go as low as the threshold of hearing (9) and for technical devices until a specified noise floor is reached. Both the upper and lower limit produce a dynamic range that can be exploited in an acoustic path culling method which can be adjusted to meet the application requirements. For example, in Virtual Reality applications, these adjustments may be pushed to a minimum to accelerate simulation computation by cutting down input data, while for expert listening tests another setup is required.

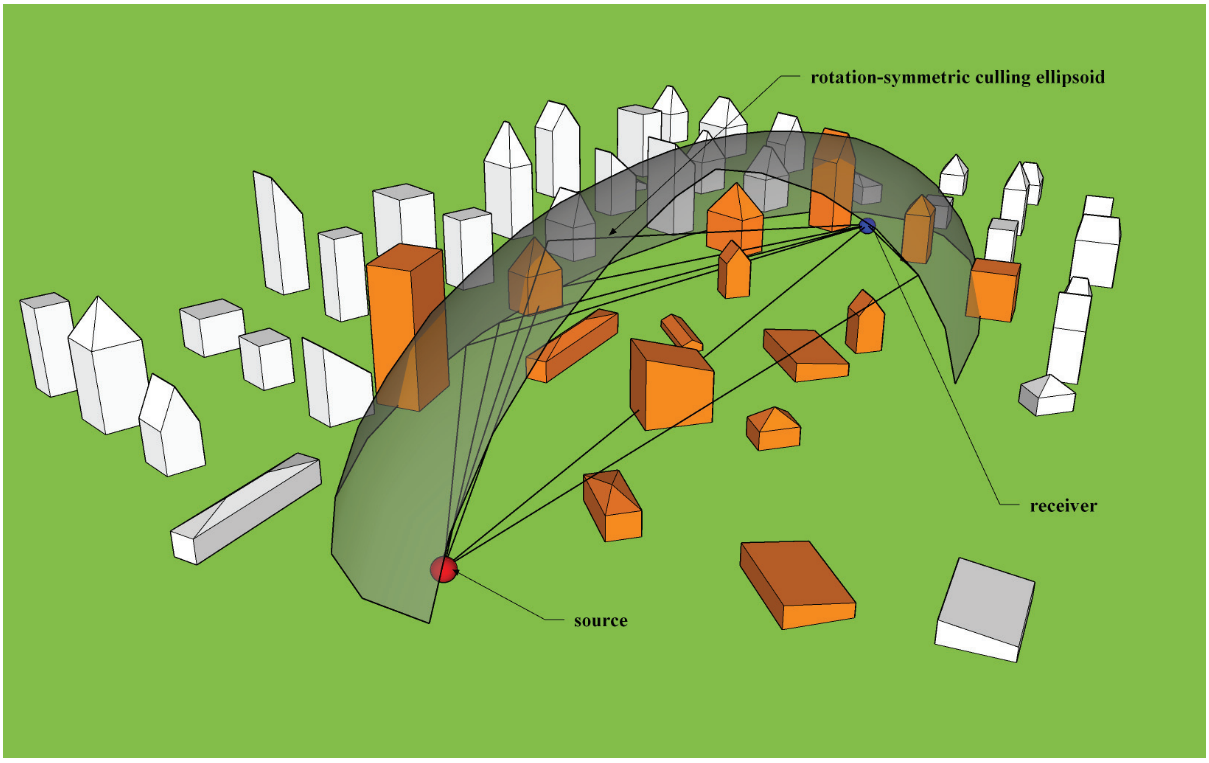

Figure 1: Culling ellipsoid for a source-receiver-pair in an urban environment.

3.2 Distance-dependent path culling

The fundamental concept of the presented approach is the assumption, that sound waves, if examined separately along the wave front normal, are solely subject to attenuation of various kinds. The most prominent influence can be seen in the spreading loss or geometrical diversion, but also reflection, diffraction and air absorption play a role. After emitted by the source, amplitudes are monotonously fading in the entire frequency range. Predominantly the impact of spherical spreading loss can be regarded as the most natural aspect of spanning a spatial area of relevance around the sound source. Fortunately, virtual acoustics commonly renders scenes where the assumption of a point source can be made, and the total sound field at a receiving point is resulting by coherent superposition of all individually determined propagation paths. Hence, if each path is separately investigated and a distance-dependent culling algorithm is applied, regions where constructive interference of sound waves occurs may well possibly lead to increased levels at greater radii from the source, e.g. when reflections focus energy. However, when searching for propagation paths, the combination of source and target positions pose the initial task: which of the surrounding mesh items must be integrated for the path construction? Figure 1 shows an example of a simple virtual urban environment where a few sound path candidates are depicted, that include a first order reflection or diffraction. The paths indicate a *detour* over faces and edges that becomes larger the farther an object is located from the axis between source and receiver. For far distances, the spreading loss itself will largely dampen the transmitted sound and is likely to undergo perceptual thresholds, either dictated by the auditory system as ultimate rational or by a context-dependent defined threshold that is adjusted by applicability. In both cases, an estimate of the maximum sound power level of the source must be known a-priori. In contrast, the definition of a fixed dynamic range reserved for acoustic transmission may be a feasible configuration if the sound characteristics can be assumed to lie in a similar region. Once the remaining dynamic range – or maximum distance - is exhausted, the propagation path search can be aborted. The determination of the maximum allowed distance for the culling algorithm is derived from the ISO formula of sound pressure level L_p at receiver given a point source with sound power L_w and the distance-dependent spherical spreading loss

$$L_p(d) = L_w - 20 \log_{10} \frac{d}{d_0} - 11 \text{ dB} \tag{1}$$

with $d_0 = 1$ m. Converted to obtain the maximum distance for an adjustable perception threshold L_{thresh}, it follows

$$d_{\max} = d_0 \cdot 10^{\frac{L_w - L_{\text{thresh}} - 11}{20}} \tag{2}$$

To give a short example, a distance of about 300 m would be found for a sound power of 90 dB and a threshold level of 30 dB without any distinct refinement of the precondition. For simplicity, no frequency dependent solution is presented with the comprehension that lower frequencies are not damped extensively over distance, except for the geometrical diversion term, though can be largely perceived by a human. Because searching paths in geometrical acoustics is a broadband approach and only gains from frequency-separate processing if a level-of-detail concept is applied, worst-case conditions must be assumed. This limitation would lead to a benefit of frequency-dependent considerations only in rather unusual situations, for example, when a source with exceptionally pronounced high-frequency components is marked relevant, although the air attenuation during transmission would largely attenuate the source's spectrum.

The region of interest reflected by the *constraint of a maximum detour* is constructed with triangles represented by vertex points of the source, the receiver and the contact point at the boundary. In three-dimensional space, the mathematical formulation results in a rotation-symmetric ellipsoidal hull that practically exhibits focus points at source and receiver, which is merely a property inherited by the rotation symmetry, but otherwise not found in ellipsoids. The reduction of the problem to a two-dimensional setup is depicted in Figure 2 and reveals this relation.

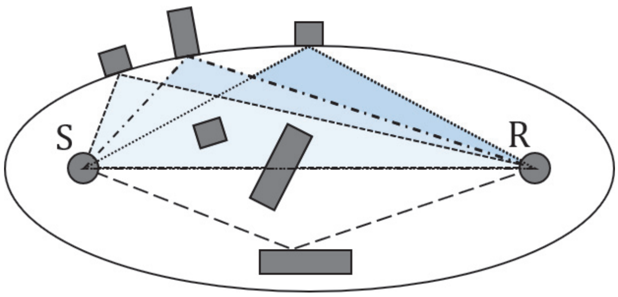

Figure 2: Cross-section of the propagation constraint of maximum detour reduced to a 2D ellipse problem.

A transform from the global into the local coordinate system of the ellipsoid results in the simplified conditional term for an internal point test following

$$\frac{x^2}{a^2} + \frac{y^2 + z^2}{b^2} < 1 \tag{3}$$

with a, b the major and minor axes of the ellipse. For the moment, the transformation must be applied to all vertex points of the building mesh under test. Transferred to a typical urban situation, a virtual hull for a given source-receiver pair is depicted in Figure 1 indicating an inner and outer region separating buildings that are relevant from those considered irrelevant, respectively. Buildings lying within the ellipsoidal hull are kept for the acoustic propagation path finding, the residual data is skipped. Performing such a filtering as a pre-processing step is able to simplify the geometrical input data set largely, however is only increasing efficiency if the test is lightweight and reliable.

3.3 Geometry data filtering for a rotation-symmetric ellipsoidal hull

Testing the intersection of one object with N other objects exhibits a linear complexity, if the intersection computation is constant. Hence the decision process, whether a building is inside or outside the rotation-symmetric hull, has linearly increasing runtime connected to the number of buildings, which is desirable. The faster the single intersection test can be computed, the better the overall result. An efficient algorithm performs the filtering (or *culling*) of geometrical objects by approximating the mesh items (faces, edges) of a building on *convex hulls* making an efficient three-folded intersection test possible. In the context of urban sound auralization, an appropriate solution is proposed by running a pre-test on the intersection of a *spherical hull*. It is constructed by the center point located at the barycenter of the complex building geometry with the radius determined by the farthest mesh point distance relative to the barycenter. The suggested approach requires semantic layers connecting buildings with corresponding geometry. Additionally, a pre-processing step generating the building hulls must be performed, which is moderate in computation time and remains constant for static objects. In case of data formats based on the CityGML convention, such properties are already largely at hand or can be simply extended, and even data from lower detail levels can be used, for example, from shoeboxes of LoD-1, if feasible for the given virtual situation.

The culling algorithm can be depicted from the flow diagram in Figure 3 and show the culling procedure that must be performed for every building. Firstly, the sphere hull of the building under test is used and an intersection of the outer sphere that fully contains the ellipsoid is constructed. Any object outside the outer sphere is considered irrelevant and can be skipped from further examination.

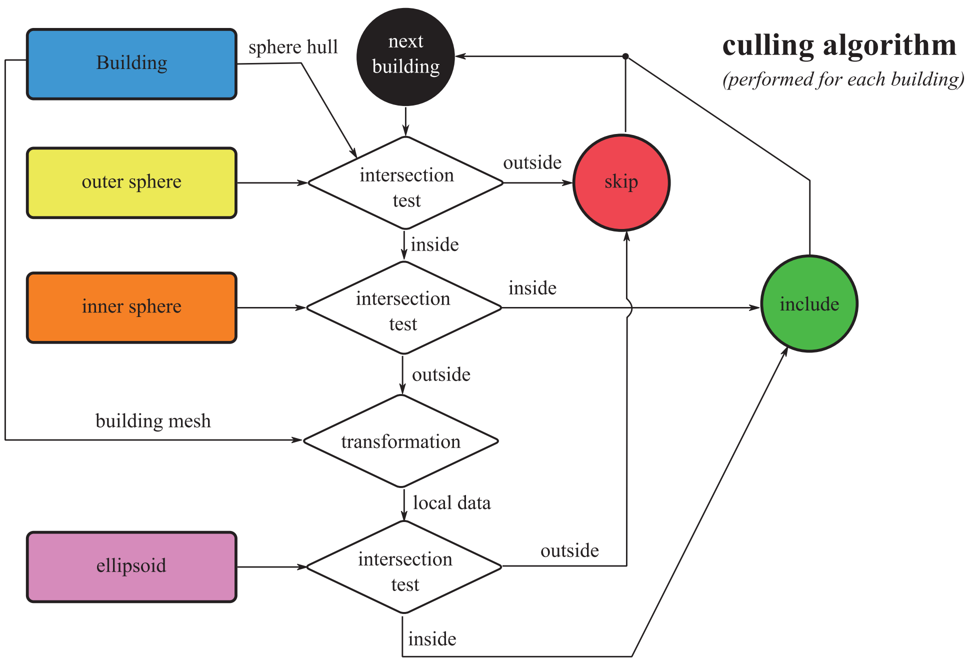

Figure 3: Culling algorithm flow diagram.

The required *intersection test* breaks down to a conceivably simple distance measure, because two spheres only intersect if the distance between the center points is smaller than the sum of the sphere radii. Another advantage is, that spheres are rotation-symmetric and hence a transformation of the building mesh is unnecessary, which makes this test remarkably fast. Accordingly, an inner sphere is constructed that lies entirely inside the ellipsoid. Any object within this inner sphere is automatically relevant and must be included. A third test is performed on the residual data set that is not marked irrelevant but still lies outside the inner sphere. In this indefinite case, a more sophisticated and costlier test is performed that evaluates the actual mesh data by transforming the mesh points into the ellipsoid reference frame. Equation (3) is then evaluated and the building is considered irrelevant, if any of the mesh points lies inside the hull, otherwise it is included. This approach inherits an approximation error that leaves objects undetected if the curvature if the ellipsoid is steep (relatively short maximum distance) and the objects under test reveal relatively large flat facades, hence precautions have to be carried out if such data may be expected. On real urban data it could be observed, that only a few percentages of buildings fall into the indefinite data set and the possibility that a falsely excluded object may have a perceivable effect, i.e. constructs a valid path that would have been audible, remains highly improbable. This is supported by the fact that close-by source-receiver pairs with a practical dynamic range for the integration of propagation paths based on the proposed algorithm construct almost spherical ellipsoids without a steep curvature in the sense that it may play a role for a failure of the test. A visualization of the culling result performed on a real data set of an inner city is shown in Figure 4, where the circles represent a sound source and a receiver, respectively, and the buildings colored in orange are marked relevant buildings for a further propagation simulation, while transparent buildings are omitted. The shape of an ellipse can be read from the image.

Figure 4: Visualization of the acoustics-driven geometry filtering algorithm applied to a real urban data set. To emphasize the resulting ellipsoidal shape, an exaggeratedly loud sound source has been used.

3.4 Extending the filter algorithm for higher reflection and diffraction orders

The presented culling algorithm has been introduced for a single reflection or diffraction component along the propagation path. Without restriction, this concept can be extended to *higher order reflections and diffractions* (including any combination). For the straightforward determination of geometrical propagation paths emitted from the source and received by the sensor, interactions follow Snell's law for reflections and Fermat's principle for diffractions. The challenge of increased complexity concerning verification tests is extremely demanding, because they have to be carried out on the entire data set. Any attempt to restrict the relevant geometry accelerates the procedure substantially. The proposed culling algorithm can be iteratively performed on the residual data set order after order, as every additional interaction of a path with the geometry concludes in an additional detour. Before a constructed propagation path is verified, for example, by an intersection test that decides if the path is audible, an integrated culling detection following the proposed algorithm can be performed that is able to sort out impossible paths efficiently. It can be expected that propagation paths departing from the direct line-of-sight between source and receiver become increasingly improbable and can therefore be removed at a very early stage of the acoustic simulation.

4. EVALUATION

Runtime evaluations have been carried out on a standard desktop computer with an Intel® Core™ i7-7700 CPU and 8 GB DDR4 memory running a Microsoft® Windows 10 operating system. For the proof of linear complexity $O(N)$ and scalability based on the number of objects N, a benchmark of the culling algorithm was conducted including a) a constant dynamic range with increasing building density and b) increasing dynamic range with constant number of buildings. For controlled conditions concerning the variation of the built environment, a randomized 3D city model (10) was used, that generates an equidistant raster of variable building shapes. Separate runtimes of the three individual validation components that must be performed for each building in worst-case revealed that a maximum of about 370 ns can be estimated (cf. Table 1).

Table 1 – Averaged runtimes of individual validation components.

Validation test	Runtime average
Outer sphere intersection	52 ns
Inner sphere intersection	53 ns
Ellipsoid intersection	265 ns

The first test varied the building number from 63 to 16,000 with some exponentially increasing intermediate points. An artificial building data set was generated that increases the equally distributed density on a square area with edge length of 1280 m. Neither buildings nor source and receiver were elevated. The receiver was located at 500 m/500 m and the sound source was located at 800 m/800 m and the maximum allowed geometrical detour resulting from a constant dynamic range was determined to 600 m. Figure 4 indicates a good agreement with a linear context.

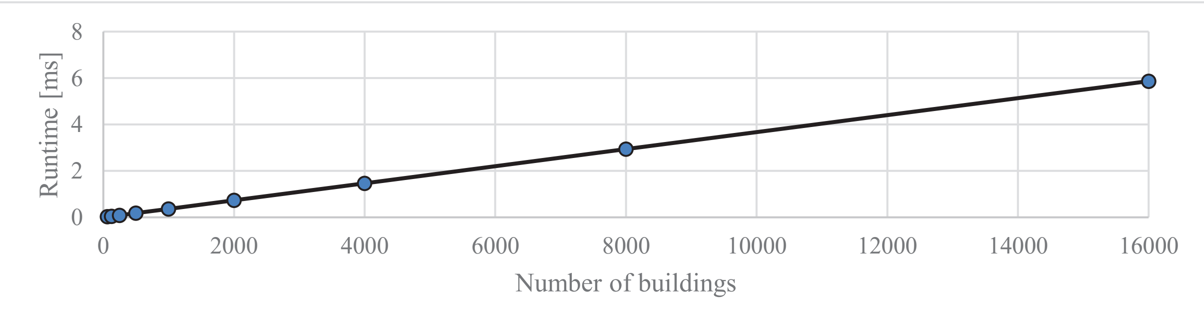

Figure 5: Runtime for increasing number of buildings at constant culling distance.

The second test distinguishes from the first test primarily by keeping the spatial extend and number of buildings constant, while varying the propagation distance between 450 m and 900 m with intermediate values at 50 m increment. Figure 5 shows only a moderate increase of runtime beginning at 2 ms for the smallest distance and linearly increasing by roughly 83 ns per 50 m increment.

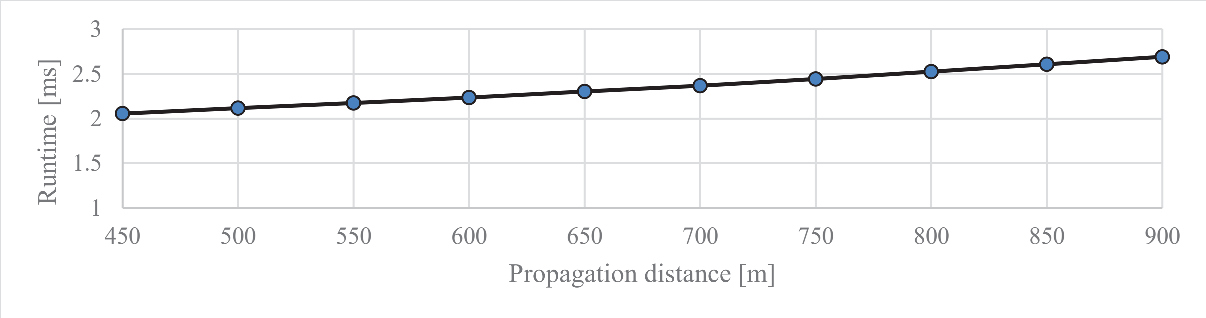

Figure 6: Runtime for increasing culling distance at constant number of buildings.

The resolution why the runtime changes over a constant building number lies in the different calculation complexity of the individual test components. With increasing distance, the number of buildings that cannot be culled immediately by the spherical intersection tests rises with the size of the ellipsoid. Hence, the relatively costly ellipsoid intersection test must be performed more often, explaining the result. In summary, it can be said that the algorithm performs as anticipated.

5. CONCLUSIONS

It has been shown that filtering geometrical input data from huge data sets based on considerations of acoustic transmission in urban environments and perceptual properties of the receiver can be implemented efficiently and with an adjustable configuration (dynamic range). The progressive nature of culling potentially undetectable geometrical propagation paths by removing building geometries outside a region of interest requires careful observation and experimental validation, if pushed to the limit. Otherwise, the proposed algorithm can facilitate urban sound auralization of single sound sources by reducing significantly the geometrical input data and, in consequence, easing calculation demands for subsequent acoustic simulations, which otherwise would lead to an impossible calculation runtime.

ACKNOWLEDGEMENTS

The authors would like to thank the German Research Foundation (DFG - Deutsche Forschungsgesellschaft) for funding the project *Auralization of Urban Environments – Real-time simulation of Diffraction* under grant number VO 600/39-1, which made this contribution possible.

REFERENCES

1. Lu L, Becker T, Löwner M-O. 3D complete traffic noise analysis based on CityGML. Advances in 3D Geoinformation. Springer; 2017. p. 265–283.
2. Kumar K, Ledoux H, Commandeur T, Stoter J. Modelling urban noise in CityGML ADE: case of the Netherlands. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Copernicus GmbH; 2017;4:73.
3. Vorländer M. Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality. Springer Berlin Heidelberg; 2011.
4. Stienen J, Vorländer M. Geometry-based diffraction auralization for real-time applications in environmental noise. The Journal of the Acoustical Society of America. ASA; 2017;
5. Kamrath M, Jean P, Maillard J, Picaut J, Langrenne C. Extending standard urban outdoor noise propagation models to complex geometries. The Journal of the Acoustical Society of America. 2018;
6. Kang J. Urban Sound Environment. Taylor & Francis; 2006.
7. Stienen J, Vorländer M. Real-time auralization of propagation paths with reflection, diffraction and Doppler shift. Fortschritte der Akustik, DAGA 2018, Munich, 2018.
8. ISO-9613-2: Acoustics – Attenuation of sound during propagation outdoors – Part 2: General method of calculation. 1st ed. International Organization for Standardization; 1986.
9. Fastl H, Zwicker E. Psychoacoustics. Springer; 2007.
10. Biljecki F, Ledoux H, Stoter J. Generation of multi-LOD 3D city models in CityGML with the procedural modelling engine Random3Dcity. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2016;