

Auralization of a car pass-by inside an urban canyon using measured impulse responses

Citation for published version (APA):

Georgiou, F., Hornikx, M., & Kohlrausch, A. (2021). Auralization of a car pass-by inside an urban canyon using measured impulse responses. *Applied Acoustics*, 183, Article 108291. <https://doi.org/10.1016/j.apacoust.2021.108291>

Document license:
CC BY

DOI:
[10.1016/j.apacoust.2021.108291](https://doi.org/10.1016/j.apacoust.2021.108291)

Document status and date:
Published: 01/12/2021

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

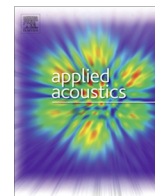
www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.



Auralization of a car pass-by inside an urban canyon using measured impulse responses

Fotis Georgiou^{a,*}, Maarten Hornikx^b, Armin Kohlrausch^c

^aEmpa, Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Acoustics/Noise Control, Dübendorf, Switzerland

^bDepartment of the Built Environment, Eindhoven University of Technology, Eindhoven, The Netherlands

^cHuman-Technology Interaction group, Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Eindhoven, The Netherlands



ARTICLE INFO

Article history:

Received 5 March 2021

Received in revised form 30 June 2021

Accepted 2 July 2021

Available online 17 July 2021

2010 MSC:

00-01

99-00

Keyword:

Auralization

Urban environments

Car pass-by

Subjective evaluation

ABSTRACT

Development of methodologies for the auralization of moving cars can be of great value for a virtual acoustic experience of urban areas. In this paper¹ a methodology for the auralization of car pass-by based on measured binaural impulse responses (BIRs) is presented. Measured BIRs for different locations in a street canyon were convolved with dry synthesized car signals in which cross-fade windows were applied in order to create a smooth transition between the source positions. Next, the convolved signals are summed in order to create the final car pass-by auralizations. A same/different listening test was carried out in order to investigate if increasing the angular spacing between the discrete source positions affects the perception of the auralizations. The experiments revealed that the auralizations with a larger angular increment (up to 16°) are perceptually different to the reference auralizations (2° spacing), even in the case where the increment increased by only 2° compared to the reference. Compared to previous listening experiments of a car pass-by in the scenario without buildings it was shown that the discrimination performance of the subjects was significantly better compared to the test conditions where buildings are absent, where subjects found it very difficult to distinguish differences between auralizations of larger and smaller increment.

© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The past decade, there has been an increased interest in the auralization of urban environments. Most of these auralization methodologies make use of engineering methods, which are broadly used in urban noise modeling. Engineering methods are based on geometrical acoustics methods, such as the image source method, where the behavior of sound waves is modeled following the principle of light rays [1]. Some popular engineering methods are Harmonoise [2], CNOSSOS [3] and Nord2000 [4]. Forssén et al. [5,6] auralized cars passing by for an environment where only the ground and a noise barrier is present. They developed a method to synthesize the car signals and they used engineering method to compute the propagation path and applied resampling to model the Doppler effect. In a later publication [7], they performed sub-

jective tests where they compared the auralizations against recordings, and approximately 50% of the subjects could not detect the differences. A method related to Forssén et al. was developed by Pieren et al. [8]. They auralized an accelerated car-pass by in an environment where only the ground reflection was present using time-variant filters that simulated the propagation delays, Doppler effect, geometrical spreading, ground reflection and air absorption. However, they did not perform subjective evaluation. Other researchers like Maillard and Jagla et al. [9,10] and Viggen et al. [11] have designed auralization methodologies for more complex outdoor environments (including the effect of buildings) with moving sources. Both methods make use of engineering methods for the computation of the sound propagation path: Maillard and Jagla et al. use the Harmonoise model [2] and Viggen et al. [11] use the Nord2000 [4]. Maillard and Jagla et al. [10] method was evaluated for the scenario where a listener is moving along bike and pedestrian paths inside a city. Subjective tests were performed where subjects were asked to judge the perception of the auralization realism on a scale from 0–10. The average score was 6.9.

Methods for the auralization of outdoor environments with moving sources have also been developed in the context of

* Corresponding author.

E-mail addresses: fotis.georgiou@empa.ch (F. Georgiou), m.c.j.hornikx@tue.nl (M. Hornikx), a.kohlrausch@tue.nl (A. Kohlrausch).

¹ Parts of this study were presented at the conference: Euronoise 2018, Crete, Greece

computer gaming using both wave-based (wave-based methods are techniques that solve the wave-equation in time (or frequency) and space [1]) and geometrical acoustics techniques [12–14]. However, for gaming applications, it is sufficient to produce plausible approximations as long as it is computationally efficient such as it fits the computational budget [1]. Moreover, none of these methods has been validated via subjective tests.

In previous work by the authors of this paper [15], auralizations of car pass by were implemented for the simplified scenarios where buildings are absent, and for an environment where a long flat wall is located behind the car, using BIRs computed with the wave-based pseudospectral time-domain method (PSTD). A dry synthesized car signal was convolved with the binaural impulse responses from the different source locations in the street, and cross-fade windows were used in order to make the transition between the source positions smooth and continuous. A same/different listening test was carried out in order to investigate if increasing the angular spacing between the discrete source positions affected the perception of the auralizations. Signal detection theory (SDT) [16] was used for the design and the analysis of the listening test. Results showed that differences exist, although they were difficult to notice. On average, 52.3% of the subjects found it difficult to impossible to spot any difference between auralizations with larger angular spacing (up to 10°) and the reference auralization (2° angular spacing).

In this paper the auralization methodology of [15] is extended for urban canyons using measured BIRs. Measured BIRs inside urban environments have not previously been reported in the literature for the auralization of moving vehicles. After post-processing the measured BIRs the same methodology as in [15] is followed to simulate the car pass-by. Finally, a same/different listening test is carried out in order to investigate if increasing the angular spacing between the discrete source positions affects the perception of the auralizations.

In Section 2 the measurement methodology and post-processing of the measured BIRs is described. In Section 3 the auralization methodology is presented. Next, in Section 4 the subjective test methodology is presented. Finally, in Section 5 the results from the subjective tests are presented and compared against the results from the listening tests presented in [15].

2. Impulse response measurements in an urban canyon

2.1. Test site

A geometrical model of the test site where the measurements took place is shown in Fig. 1. The test site is located in the outskirts of Eindhoven in an industrial area. The reasons that this test site was selected are: 1) the geometry of the canyon between Building A and the three smaller buildings (B,C,D) (all 4 buildings have heights between 11–13 m) is very similar to an urban canyon, 2) it is located away from residential areas, which allowed the source to be played at high power levels thereby achieving a higher signal-to-noise ratio, 3) on the weekends the site is closed so no pedestrians or cars passed through the street. The test site consists of logistics warehouses.

In SubSection 2.3 the measurement procedure is described together with actions taken to minimize the effect of those noise sources in the BIR measurements.

2.2. Measurement set-up and equipment

The measurement equipment is listed in Table 3. A Toshiba laptop running on Windows 8 and the Dirac 6.0 impulse response (IR) and analysis software were used for the measurements. The output

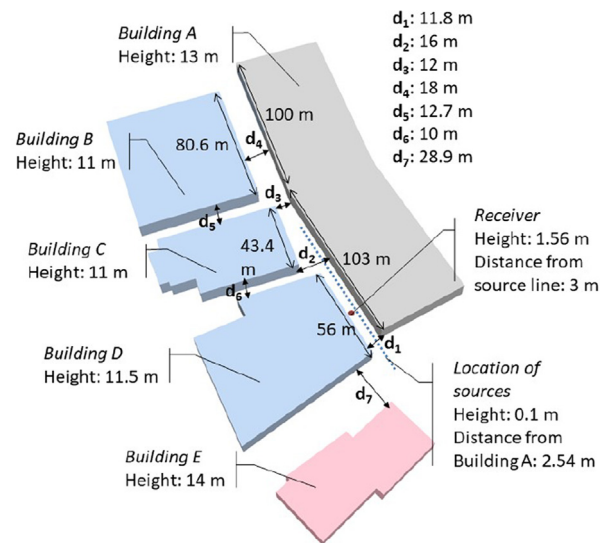


Fig. 1. Geometrical model of the test site.

of the laptop was sent to a Triton sound card via a USB cable and the output of the sound card was connected to an Amphon power amplifier. The signal from the amplifier was sent via 2×20 m cables with Speakon connectors to the dodecahedron source. The outputs from the two ear microphones of the head and torso simulator (HATS) were sent via 4 m cables with BNC connectors via the sound-card to the laptop.

At octave-bands below 1 kHz the dodecahedron source becomes almost perfectly omnidirectional, it is somewhat directional in the 1 kHz octave-band and above 1 kHz the source becomes highly directional. Specifically, there is almost a 4 dB variation in the directivity pattern of the 1 kHz octave-band [17]. More information about the directivity of the dodecahedron source can be found in Hak et al. [17].

A detailed description of the test site and the measurement plan can be found in Fig. 1. The source was placed on the ground on top of a 4 cm thick rubber panel and was not rotated during the measurements. The distance between the lowest speaker of the source and the ground was approximately 10 cm (the center of the dodecahedron source was approximately at 30 cm above the ground surface). The reason for this position is that it allowed BIRs to be captured from a near-to-ground source location in order to use them for auralizations of a car pass-by where the main noise source for speeds above 30–40 km/h comes from the tires [2,4,18]. However, this approach is not ideal and highlights a significant limitation in the auralization of moving cars, since as mentioned in [2,4,18], the main noise sources that make up the car sound (exhaust, engine, and tire noise) are located at different heights. For this reason, in engineering methods such as Harmonoise and Nord2000 [2], the car is modeled by point sources at different heights. The distance between the source and the façade of Building A (Fig. 1) was 2.54 m. The HATS was mounted on a stand facing the wall of Building A and was placed in the middle of the façade of Building D. The ears were at 1.56 m distance above the ground. The distance between the HATS and the source line was 3 m.

In order to minimize the wind effect on the in-ear microphones, windshields were mounted in the ears using tights (See Fig. 2). The windshields were B&K windshields for 1/2 inch microphones, which were cut in half. In order to fit the windshields in the HATS ears some foam was removed from the inside of the windshields to open a hole at the size of the HATS ears. To assess the effect of the windshield on the measured levels, IR measurements were



Fig. 2. B&K HATS with windshields.

performed under laboratory-controlled conditions with the HATS wearing windshields and tights and without inside the transmission room of the Echo building at TU Eindhoven Campus using the B&K Omnipower source 4292-L. The effect of the windshield and the tights is shown in Fig. 3. As demonstrated in Fig. 3, the windshield caused a drop of 1–2 dB at frequencies between 4–10 kHz and had zero effect on frequencies below 4 kHz.

2.3. Measurement procedure

A measurement grid was manually created using a tape measure and gaffer tape (see Fig. 4) to plot the distance between the points on the street (the points were located along the dashed line in Fig. 1). Since the scale is relatively large (70 m), the small errors between the points (± 5 cm) were considered reasonable. Measurements were taken for every 2° (angle between two neighboring measurement points and the receiver).

The measurements were taken on the 30/07/2016 between 06:00–12:30 when all the warehouses were closed in order to minimize the presence of disturbing background noise.

The main noise sources in the area during the measurements are listed below:

1. Traffic noise from the highway located 270 m from the site;
2. Construction noise coming from approximately 170 m away (produced behind Building E) See Fig. 1). The main noise source was drilling, although it was not constant;
3. Plane flyovers at a frequency of 6–12 per hour for a duration of 3–45 s;

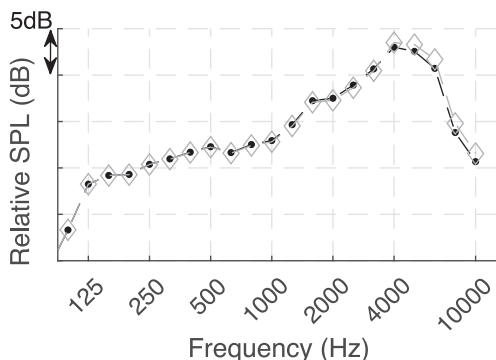


Fig. 3. 1/3 octave-band energy plots of the BIR measured with the HATS inside the transmission room of the Echo building at TU Eindhoven with (dot) and without (diamond) windshield. The BIR was not corrected for the source spectrum (SPL refers to sound pressure level).

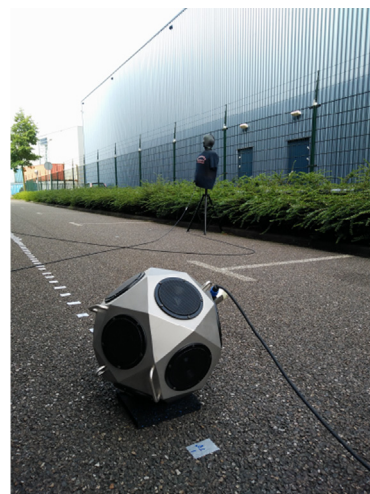


Fig. 4. Measurement photo. The source was placed on top of a 4 cm thick rubber panel. The gaffer tape was used to create the measurement grid.

4. Intermittent wind noise;
5. Intermittent loading and unloading of the trucks from building E (see Fig.1).

Measurements were not taken at times when: planes flew over; a vehicle passed by at a near distance; construction work was considered to be noisy. The average wind speed during the measurements was approximately 0.7 m/s and the maximum speed was 3 m/s. The wind was measured at a distance of 5 m from the HATS and at 1.6 m above the ground. The measurement was repeated if the wind exceeded the speed of 3 m/s. The average temperature was 20 C° and average humidity was 70.5%.

The main difficulty with outdoor measurements for auralization purposes is that it is very difficult to get good signal-to-noise ratios, especially at far distances. The time varying background noise level affects the quality of the measurements. Also, a time-varying condition caused by wind can alter the transfer function between the source and the receiver. In order to avoid risking changes in the weather and to ensure enough battery life during the whole duration, the measurements had to be taken efficiently.

Multiple trial measurements were implemented in order to determine which measurement signal gave the best results. The distance between source and receiver in those trial measurements was 30 m. Both maximum length sequence (MLS) and exponential sine sweep (e-sweep) test signals [19] were used. The e-sweep clearly achieved the best impulse-response-to-noise ratio (INR)². Long e-sweeps did not produce a high INR because during the measurements the background noise and wind speed varied. The best results were produced by the shorter e-sweeps which were repeated multiple times in order to improve the INR. As such, it was decided to use e-sweeps of 3×10.9 s. In Fig. 5 the INRs (computed with the Dirac 6.0 software) of 3 measured BIRs at different locations are plotted. At low frequencies the INR dropped significantly with increasing source-receiver distance. The INR reduced less at frequencies above 1 kHz as distance increased. The main reason for this finding is the low frequency nature of the background noise sources in the measurement environment.

The most distant BIR used in the auralizations of the car pass-by was the one measured at $\pm 82^\circ$ (0° angle is at the middle of the measurement grid where left and right ear BIRs were assumed to

² INR is equal to the maximum root-mean-square (RMS) level of the IR (in dB) minus the noise level (in dB)[20]

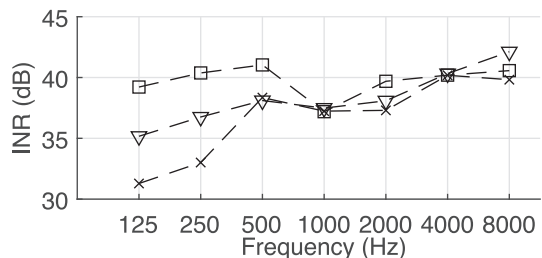


Fig. 5. Impulse response to noise ratio of the right ear BIRs measured along the receiver line of Fig. 1 for different octave-bands at measurement angles 8° (square), 16° (triangle), and 82° (cross). These locations correspond to a source/receiver distance of 5.2 m, 10.46 m and 21.55 m. 0° angle is at the middle of the measurement grid, -82° angle is at the left side and 82° angle at the right side of the HATS.

be symmetrical, -82° angle is at the left side and 82° angle at the right side of the HATS. In Fig. 6 a 2D presentation of the BIRs measured in the left ear’s channel at ±82° between buildings A and D (see Fig. 1) is shown.

2.4. Post-processing of BIR measurements

A high frequency artefact was present in the right channel (ear) measurements due to a defect in a measurement cable. In order to solve this issue, the left ear BIRs were mirrored for the auralizations of the right ear. For example, the right ear BIR measured at 82° was replaced with the left ear BIR at -82° and the right ear BIR measured at -82° was replaced with the left ear BIR at 82°. This mirroring was possible since from the position of 28 m to the left and right sides of the HATS the street canyon is almost symmetrical.

The frequency response of the dodecahedron source is not flat, so the measurements had to be corrected. In Fig. 7 the average spectrum of the dodecahedron source measured inside an anechoic chamber at 45° intervals is plotted. It can be seen in this figure that the dodecahedron source does not generate much energy at frequencies below 50 Hz and above 9 kHz. The spectrum of the source was corrected from the BIRs for frequencies between 50 Hz - 9 kHz by designing an inverse filter based on the response shown in Fig. 7. The inverse filter was designed using the least-square decon-

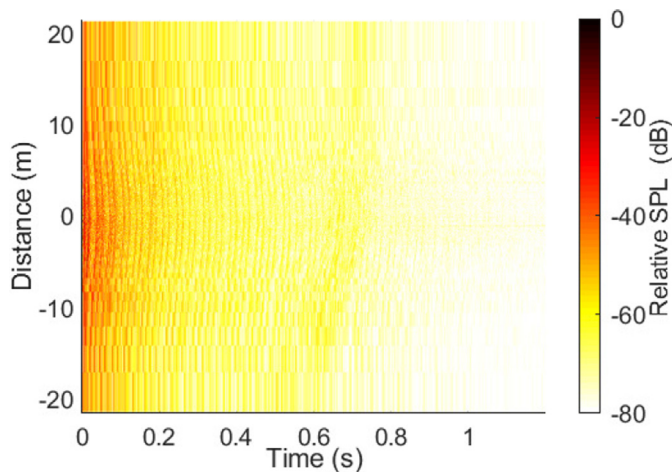


Fig. 6. 2D sound field presentation of the urban canyon between buildings A and D using the BIRs measured ±82° from the HATS (see Fig. 1). The measured BIRs of the left ear were normalized based on the maximum amplitude from all BIRs and their initial delay was removed in order to be time aligned. The colour bar shows the SPL of the normalized left ear BIRs in dB.

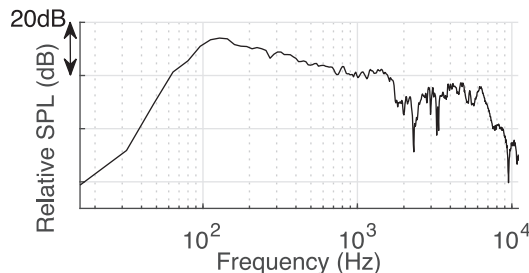


Fig. 7. Average frequency response of the dodecahedron source. The source was measured inside an anechoic chamber by rotating every 45° (8 measurements in total) and the average spectrum is plotted here.

volution method with frequency-dependent regularization as presented in Section 2.5 in [15]. Fig. 8 presents a measured BIR (both time signal and spectrum) before and after the inverse filter was applied to correct the source spectrum.

3. Auralization methodology

The dry car signals used in this work (see Fig. 9) were provided by Chalmers University of Technology and are synthesized with the Listen project simulator [5]. The synthesized car signal was split at increments based on the time it takes the car to travel between 3 discrete source locations. A sine window function was applied to these signals in order to achieve a smooth cross-fade between the signal from the source positions. Next, these signals were convolved with the corresponding measured BIRs and shifted by the time that it takes the car to travel between two discrete neighbouring source positions. Finally, these signals were added together to create the final binaural auralization stimuli. The details regarding the synthesized car signals, the auralization methodology and its limitations can be found in [15].

The Doppler effect was not included in the auralizations. The Doppler effect is a dynamic effect and cannot be modeled by switching between static sources as in the work presented in this paper. Thus, further processing, is required to simulate the Doppler effect. Simplified methods, which potential applicability in this

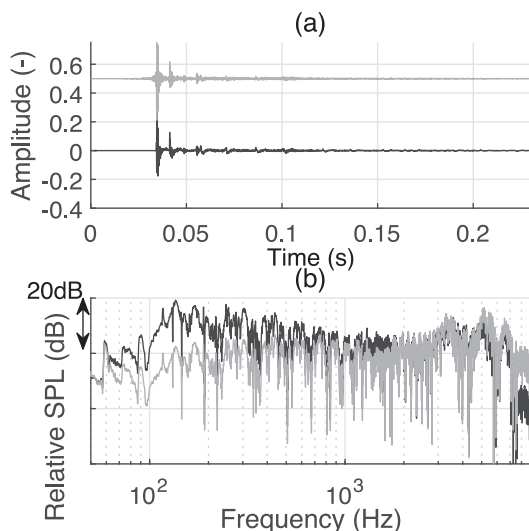


Fig. 8. Measured right ear BIR at 74° along the receiver line (see Fig. 1) before (black) and after (grey) correcting the spectrum of the dodecahedron source (see Fig. 7). In (a) the time response is truncated at 0.2 s (whole duration is 2.5 s) and a 0.45 offset was added to the corrected BIR to make both lines visible. In (b) the frequency response of the full BIR is plotted.

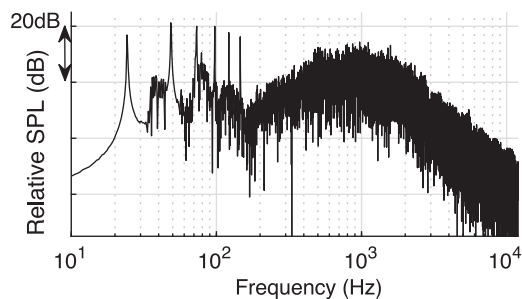


Fig. 9. Spectrum of a dry car signal (70 km/h) with tonal components used in the auralizations presented in this work.

methodology, have been developed to model the Doppler effect using variable delay lines and interpolation of the signal (asynchronous resampling process) [8,21]. The main focus of the listening tests presented in this paper was the perception of the switch between the discrete source positions. Therefore, it was decided to exclude the Doppler effect from the listening tests stimuli and focus entirely on the switch between the discrete source positions. The impact of the Doppler effect on the perception of the switch between the discrete source positions when the source increment is increased was left for future work.

4. Subjective test

4.1. Test methodology

Because large source spacing allows for fewer IRs measurements, less computational storage requirements and fewer convolutions for the creation of auralizations, it is important to investigate how large the spacing between the BIRs can be without affecting the perception of the car pass-by auralization [22]. A same/different test [23] was conducted in which the subjects were presented with two stimuli and had to indicate whether the stimuli were the same or different. The first stimuli was the auralization of a car pass-by inside the street canyon with the reference angular spacing (2° angle between two neighbouring sources and the receiver) and the second the auralizations with larger angular spacing. The reason that a same/different test was chosen is that the task is subject-friendly because the responses "same" and "different" are familiar and easily understandable [24]. The design and the analysis of the test was based on signal detection theory (SDT). The basics of SDT that were used in this research can be found in [15] and Section 4.4 of Georgiou [22]. The reason that the 2° increment was selected as reference increment is that it was the smallest angular increment that produced a smooth and continuous cross-fade for cars moving at speeds of 50 km/h and 70 km/h.

The same test methodology presented in [15] was followed; the same design, test interface (Max 7 software), trial repetitions, playback levels (84 dB(A) maximum playback level), equipment (Sennheiser HD 800 headphones, which were connected to a laptop via the E-MU 0204 USB soundcard). Each of the 3 test conditions presented in Table 1 consisted of 50 trials; 25 of those trials were reference signal versus reference signal and the remaining 25 reference signal versus a signal auralized using a larger angular increment. Each test condition was tested individually; i.e. subjects had to complete 50 trials of that condition in one run. For every subject the test conditions order and the trials of each condition were randomized. Prior to starting the test, the subjects were given written and oral instructions (no instructions were given on the kind of differences that subjects should pay attention to).

The listening test from [15] showed that different speeds and car signals (with and without tonal components) did not produce

Table 1

List of the 3 different test conditions and their notation.

Speed (km/h)	Tonal components	Increments tested vs 2°	Notation
70	Yes	4	70TM4
70	Yes	8	70TM8
70	Yes	16	70TM16

different results. In this experiment only one vehicle speed was selected (70 km/h) in order to shorten the duration of the test. The same car signal with tonal components as in [15] was chosen because: 1) real car signals include tonal components; 2) cross-fading between the car signals blocks that include tonal components is more likely to produce audible changes when the increment between the discrete source positions is increased compared to the auralization with car signals without tonal components (see Section 3.3 in [15]). Thus, if subjects find it hard to perceive differences between the reference auralization and auralizations with larger increments for car signals with tonal components, this would also stand for the case without tonal components. As in the work presented in [15], no background noise, which is an important part of the urban soundscape [25], was added to the auralizations.

The tested angular increments that were compared against the reference are 4° , 8° , 16° . The reason for this choice is that in the previous experiments [15] for the simplified environments where buildings are absent and for an environment where a long building block is located behind the car, most of the subjects found it very difficult to perceive any difference between the reference auralization and the auralization with larger increments, even for angles up to 8° and 10° . Also, there were no significant differences between the different increments (subjects were equally sensitive in detecting differences in the auralizations with the 4° increment and the 8° and 10°). Therefore, it was decided to not test many intermediate increments. Instead, a larger spacing was tested, which was expected to be identified from the reference auralization by the subjects.

Finally, the subjects were asked at the end of the test to mark on a verbal scale how difficult they found the listening test. They were also asked to note the number of breaks they had during the test, and if they felt tired.

4.2. Test subjects

Fourteen subjects participated in the listening test (9 male and 5 female). Ten of those subjects had participated in the previous experiment [15]. The average age of the participants was 32.4 years. Twelve of those subjects have a profession related to sound and acoustics. Non of the subjects reported hearing problems. Only three of the subjects found the playback level comfortable (maximum instantaneous SPL of 84 dB(A)). The maximum playback level for most subjects was adjusted to 81 dB(A) (for two subjects this value was set to 78 dB(A)). The duration it took each subject to complete the test was approximately 20–25 min.

5. Results and comparison with the previous listening tests

5.1. Results

A fundamental measure of difference in the perception of the two stimuli is the sensitivity or d' value. The sensitivity d' is a measure of the subjects' ability to identify whether the stimuli are the same or different; the greater the d' , the easier it is for the subjects to recognize the differences. Here, the limit to consider two auralizations to be different is $d' \geq 1$, which is a limit commonly used in

the literature [26]. The computation of the sensitivity measure d' for the same/different test and the analysis of the test results was based on the method and equations used in the previous listening experiment described in Section 5 in [15]. In Fig. 10 the scores for the different test conditions are plotted in Box-and-Whisker plots. In Fig. 11 the scores of every individual for the 3 different test conditions are plotted with different markers.

One subject was excluded from the analysis because (s) he scored d' of 0 in all conditions, and as such was considered an outlier. 54% of the subjects reported that they felt tired during the test. The low percentage of $d' < 1$ shows that the vast majority of subjects could perceive the differences between the reference auralizations and auralizations with larger angular increments (even for the 4° angular increment).

In Fig. 10 and Fig. 11 it can be seen that subjects tended to better identify differences (score higher d') between the auralizations when the angular increment is increased; i.e. more subjects scored $d' > 1$ in the condition with angular increment of 8° than at 4°, and the same applies for the increment of 16° compared to 8°. Paired t-tests and Kruskal–Wallis tests were conducted between the test conditions of different angular increments and the results are shown in Table 2. The reason that Kruskal–Wallis tests were performed for some pairs is that the Kolmogorov–Smirnov test for normality gave a $p < 0.05$, so the assumption for the t-test was not satisfied. Kruskal–Wallis is a non-parametric analysis of vari-

Table 2

Level of significance between the mean d' different angular increment test conditions. p-values were computed with paired t-tests and the Kruskal–Wallis test between the d' scored on the test conditions of the left column. The asterisk indicates that the p-values have been computed with the Kruskal–Wallis due to failing the normality test.

Test condition pairs	Level of significance (p-value)
70TM4 vs 70TM8	0.7
70TM4 vs 70TM16 *	0.022
70TM8 vs 70TM16 *	0.061

ance and it was used because the data are not normally distributed [27]. The p-values of the pairs shown in the left column of Table 2 that have an asterisk were computed with the Kruskal–Wallis test. The difference is significant ($p < 0.05$) only between 70TM4 vs 70TM16.

The vast majority of subjects could perceive the difference in the auralizations when the increment was increased. However, on average subjects found the overall task of the listening test difficult. In Fig. 12 the responses of the subjects to the question regarding the difficulty of the task is plotted (subjects were asked to rate the difficulty on a scale as shown on the y-axis of Fig. 12). The mean response lies between “Somewhat difficult” and “Difficult”, with two subjects responding “Somewhat easy” and two “Very difficult”.

5.2. Comparison with the previous listening test

Here, the same experiment was conducted as that presented in [15], where an experiment was conducted using simplified scenarios in which buildings are absent and for an environment where a long wall is located behind the car. The BIRs in [15] were computed using the wave-based PSTD method. In the listening test presented in this paper the car was moving at 70 km/h and the car signal used included tonal components (the dry car signal was the same as in [15]). The same auralization cases were tested in the listening test presented in [15] for the condition where buildings are absent. These results are plotted in Fig. 13 against the results obtained from the listening tests presented in this paper in order to compare the performance of the subjects in these two conditions. It should be noted again that in both tests 13 subjects participated and that background noise was not included in the auralizations. From these subjects, 10 participated in both tests.

Subjects could distinguish differences between the auralizations with larger increment and the reference auralizations much more easily in the current street canyon case. Another outcome by comparing these results is that in contrast with the street canyon, the d' values of the simplified case did not increase when the angular increment was doubled (at 8°). The angular increment of 16° was not tested in the simplified case. Thus, it is not part of this comparison.

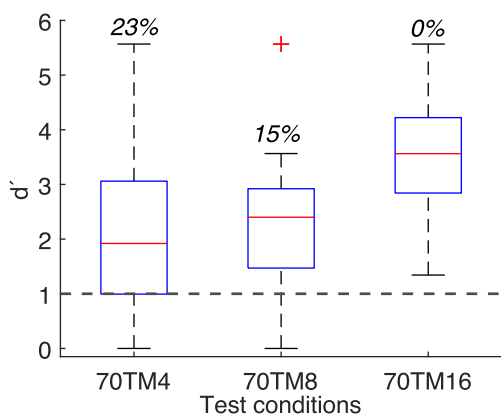


Fig. 10. Box-and-Whisker plots of the d' scored for every test condition (x-axis) e.g. 70TMx: Car moving at 70 km/h, with tonal components, auralized with measured IRs and x° angular spacing between the discrete source positions. See Table 1 for all notations. The plot shows the lower and upper quartile values, and the median value (red line). The whiskers represent the remainder of the data. The numbers on top of the plots show the percentage of subjects who scored $d' < 1$.

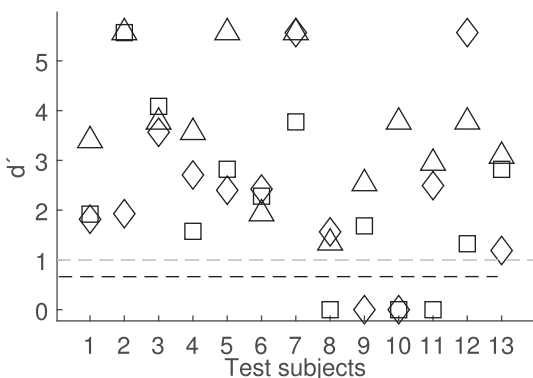


Fig. 11. Plot of the d' scored by each subject for the 3 different test conditions. The markers represent the angular increments. Square: 4°, Diamond: 8°, Triangle: 16°.

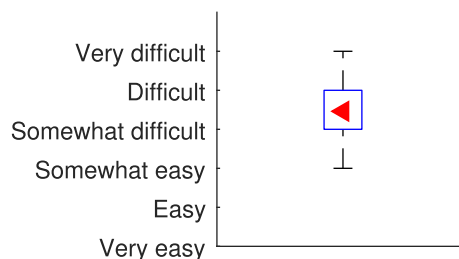


Fig. 12. Box-plot with the choice of the subjects regarding the difficulty of the listening test. The y-axis labels are the discrete responses that subjects were asked to choose in order to indicate the difficulty of the task. The red triangle indicates the mean value.

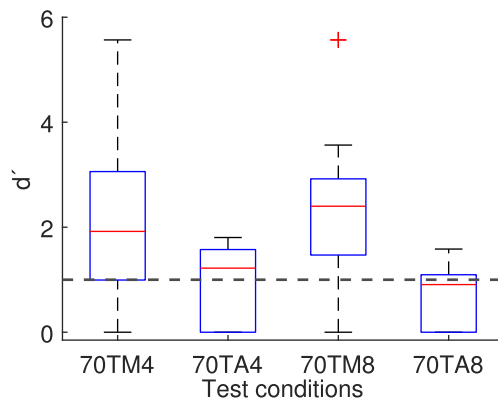


Fig. 13. Box-and-Whisker plots of the d' scored in two conditions (x-axis) in this listening test and two conditions from the listening test presented in [15] for the environment where buildings are absent. 70TMx: Car moving at 70 km/h, with tonal components, auralized with measured BIRs inside the urban canyon and α° angular spacing between the discrete source positions; 70TAx: Car moving at 70 km/h, with tonal components, auralized with simulated BIRs of an environment where buildings are absent [15] and α° angular spacing between the discrete source positions. Parts of this data are also plotted in Fig. 7 in [15] and Fig. 10. in this paper.

Experiments that look into the nature of difference and into the cues that subjects used to make their judgments could possibly provide further insight into the reasons that subjects were more sensitive to changes in the angular increment in a reverberant environment, such as the street canyon compared to the simplified case where buildings are absent. Some possible reasons for the differences between the two tests are listed below:

- Early reflections and reverberation, which were not present in the auralization of the previous experiment, could have provided more cues to the subjects to judge the differences
- Cross-fading between discrete source positions with 2 and 8 degree increments could be sufficient for an artefact free propagation delay interpolation for simple geometries, where a single path dominates the response, but not for street canyon geometries, where additional paths with independently varying delays are present.
- The duration of the street canyon test was half that of the first test. This might have affected the performance since subjects might have kept a clearer mind during the whole duration of the second test.
- A very large increment (16°) was tested in the street canyon case where almost 100% of the subjects scored $d' > 1$. This means that 1 out of 3 test conditions was fairly easy for the subjects to identify the difference. These positive identifications may have motivated the subjects and increased their focus. In the listening test presented in [15] the maximum tested angular increment was 8° and there was no easy test condition for the subjects. The difficult test conditions may have negatively affected the confidence and subsequent motivation of the subjects.
- The spectrum of the auralizations in the listening tests of paper [15] was truncated at 7.5 kHz and in this test the auralizations in the street canyon with the measured BIRs were truncated at 9 kHz. However, this fact probably did not play any role because the dry car signal used in the auralization (see Fig. 1 in [15]) did not contain a lot of energy above 6 kHz.
- The mirroring of the BIRs due to the issue on right channel measurements might have played a role on the results. However, it is more likely that it affected the realism of the auralizations, because in reality nothing is 100% symmetric as it was assumed in the research presented in this work in order to perform the mirroring.

The street canyon tests were performed 5 months after the first test, so it is highly unlikely that a training effect played a role; i.e. that the 10 subjects who participated in the first test performed better in the second one because they were trained.

6. Conclusions

This paper presented a methodology for the auralization of car pass-by inside an urban canyon using measured BIRs. This paper is an extension of [15] which is focused on simplified acoustic spaces where buildings are absent and for an environment where a long building wall is located behind the moving car. This is the first time that measured BIRs are used for the auralization of car pass-by.

The BIR measurements were conducted inside a street canyon for every 2° increments along a straight line whose distance from the receiver was 3 m. The B&K HATS used for the measurements was placed at an ear height of 1.56 m and the dodecahedron source at 0.1 m from the ground. The spectrum of the source was corrected from the measured BIRs using an inverse filter based on the average spectrum of 7 IRs of the dodecahedron source measured at rotations of 45° inside an anechoic chamber. Next, the post-processed measured BIRs for different discrete measurement locations inside the street canyon were convolved with the dry synthesized car signal and cross-fade sine windows are used in order to create a smooth transition between the source positions. A detailed explanation of the methodology is included in [15].

A same/different test was conducted to evaluate whether subjects could perceive differences between the auralization with the reference angular spacing (2°) and the auralizations with larger angular spacing ($4^\circ, 8^\circ, 16^\circ$). The experiments revealed that the auralizations with a larger increment are not perceptually identical to the reference auralizations, not even in the case where the increment increased by only 2° compared to the reference. The discrimination performance of the subjects was significantly better compared to the test conditions that were evaluated in [15] for the simplified environments.

The fact that the auralizations with larger angular increments were distinguishable does not necessarily mean that they have no use in future applications. It would be interesting to study for what kind of applications this perceptual difference is relevant e.g. is this difference important when conducting experiments on annoyance from traffic noise or experiments on the evaluation of urban soundscapes? Also, the plausibility or authenticity of the auralizations with larger increments compared to the auralizations with smaller source increments should also be investigated. It would be also interesting to perform the same listening tests but for more realistic traffic scenarios where more cars are passing by: it is possible that an environment with multiple sound sources might have an influence on the perception of the auralizations with larger increments. One of the limitations of this research is that the car was considered as a source that radiates sound from only one point (BIRs were measured 30 cm above the ground surface). Measuring BIRs at different heights associated with the main radiation sources of common vehicle noise components (e.g., exhaust, engine and tire noise) and investigating its effect on the perception could be a valuable future extension to the present work. Finally, the perceptual tests performed for this research were not focused on assessing the realism of the proposed auralization methodology. Assessing the method's realism is something that needs to be done in the future.

Finally, it should be mentioned that obtaining high quality BIRs for auralization purposes inside urban canyons is very challenging because the measurement conditions are not stable. The measurements are very sensitive to external sound sources, such as moving cars and construction noise, and changes in the weather conditions

make things extra difficult. Moreover, the measurement source needs to have more power compared to indoor environments because the measurement distances are longer and energy is lost due to the open sides of such geometry. Thus, acoustic modeling is a significantly more convenient way to obtain BIRs of urban canyons. However, the modeling method needs to be able to model accurately and relatively efficiently both low and high frequencies. Therefore, developing a hybrid wave-based and geometrical acoustics method could be of great value for the auralization of urban environments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Jan de Rijk logistics for allowing us to perform the measurements at their site.

Appendix A

Table 3
Measurement equipment.

Equipment type	Details
Laptop	Toshiba satellite with Windows 8
Measurement software	Acoustics Engineering-Dirac 6.0
Battery	Goal Zero, model Yeti 400
Dodecahedron source	B&K Omnipower source 4292-L
Power amplifier	Acoustics Engineering-Amphion
Sound card	Acoustics Engineering-Triton
HATS	B&K 4128-C
Humidity meter	Rotronic hydrolog HL-10
Wind instrument	Windmate WM-200
Cables	4 × 50 m BNC, 2 × 20 m Speakon cables

References

- [1] Välimäki V, Parker JD, Savioja L, Smith JO, Abel JS. Fifty years of artificial reverberation. *IEEE Trans Audio, Speech, Language Process* 2012;20(5):1421–48.
- [2] H. Jonasson, U. Sandberg, G. van Blokland, J. Ejsmont, G. Watts, M. Luminari, Source modelling of road vehicles, Deliverable 9 of the Harmonoise Project (2004) 1–52..
- [3] S. Kephelopoulos, M. Paviotti, F. Anfosso-Lédée, Common noise assessment methods in Europe (CNOSSOS-EU), European Commission Directorate-General Joint Research Centre..
- [4] H. Jonasson, Acoustic source modelling of Nordic road vehicles, SP Rapport 12..
- [5] Forssén J, Kaczmarek T, Lundén P, Nilsson ME, Alvarsson J. Auralization of traffic noise within the LISTEN project - Preliminary results for passenger car pass-by. In: *Proceedings of Euronoise Edinburgh, UK*.
- [6] A. Peplow, J. Forssén, P. Lundén, M.E. Nilsson, Exterior auralization of traffic noise within the LISTEN project, in: *Proceedings of Forum Acusticum, Aalborg, Denmark, 2011*, pp. 665–669..
- [7] Nilsson ME, Rådsten-Ekman M, Alvarsson J, Lundén P, Forssén J. Perceptual validation of auralized road traffic noise. In: *40th International Congress and Exposition on Noise Control Engineering, Osaka, Japan*. p. 3453–60.
- [8] Pieren R, Büttler T, Heutschi K. Auralization of accelerating passenger cars using spectral modeling synthesis. *Appl Sci* 2016;5(6):1–27.
- [9] J. Maillard, J. Jagla, Auralization of non-stationary traffic noise using sample based synthesis - Comparison with pass-by recordings, in: *Proceedings of Internoise, Saint-Martin-d'Hères, 2012*..
- [10] J. Maillard, A. Kacem, Auralization applied to the evaluation of pedestrian and bike paths in urban environments, in: *Proceedings of Internoise, Hamburg, Germany, 2016*, pp. 3473–3482..
- [11] Vigen EM, Audun S, Venerød J, Olsen H. Development of an outdoor auralisation prototype with 3D sound reproduction. In: *Proceedings of the 18th International Conference on Digital Audio Effects, Trondheim, Norway*.
- [12] R. Mehra, N. Raghuvanshi, L. Antani, A. Chandak, S. Curtis, D. Manocha, Wave-based sound propagation in large open scenes using an equivalent source formulation, *ACM Transactions on Graphics* 32 (2) (2013) 19:1–19:13..
- [13] N. Raghuvanshi, J. Snyder, R. Mehra, M. Lin, N. Govindaraju, Precomputed wave simulation for real-time sound propagation of dynamic sources in complex scenes, *ACM Transactions on Graphics* 29 (4) (2010) 68:1–68:11..
- [14] H. Yeh, R. Mehra, Z. Ren, L. Antani, Wave-ray coupling for interactive sound propagation in large complex scenes, *ACM Transactions on Graphics* 32 (6) (2013) 165:1–165:11..
- [15] Georgiou F, Hornikx M, Kohlrausch A. Auralization of a Car Pass-By Using Impulse Responses Computed with a Wave-Based Method. *Acta Acustica united with Acustica* 2019;105(2):381–91.
- [16] Macmillan NA, Creelman CD. *Detection theory: A user's guide*. Psychology press; 2004.
- [17] C. Hak, R.H.C. Wenmaekers, J.P.M. Hak, L.C.J. van Luxemburg, The source directivity of a dodecahedron sound source determined by stepwise rotation, in: *Proceedings of Forum Acusticum, Aalborg, Denmark, 2011*, pp. 1875–1879..
- [18] Huijssen J, Hallez R, Pluyms B, Desmet W. A synthesis procedure for pass-by noise of automotive vehicles employing numerically evaluated source-receiver transfer functions. *J Sound Vib* 2013;332(15):3790–802.
- [19] Farina A. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In: *Proceedings of the Audio Engineering Society 108th Convention, Paris, France*.
- [20] . Tech. rep. 2002.
- [21] H. Strauss, Implementing doppler shifts for virtual auditory environments, in: *Audio Engineering Society Convention 104*, 1998..
- [22] F. Georgiou, Modeling for auralization of urban environments: Incorporation of directivity in sound propagation and analysis of a framework for auralizing a car pass-by, Ph.D. thesis, Eindhoven University of Technology (2018)..
- [23] Lee HS, van Hout D, Hautus M, O'Mahony M. Can the same-different test use a β -criterion as well as a τ -criterion? *Food Qual Prefer* 2007;18(4):605–13.
- [24] Liang J, Zhou Y, Liu Z. Examining the standard model of signal detection theory in motion discrimination. *J Vision* 2016;16(7):1–12.
- [25] Zachos G, Forssén J, Kropp W, Mauriz LE. *Background traffic noise synthesis 2016*.
- [26] Le S, Worch T. *Analyzing sensory data with R*. Chapman and Hall/CRC; 2014.
- [27] Hollander M, Wolfe DA. *Nonparametric statistical methods*. 2nd Edition. New Jersey: John Wiley & Sons Inc; 1999.