Capturing the External Sound Sources of Trains Using Beamforming

Fanyu Meng, Frank Wefers, Michael Vorlaender

Institut für Technische Akustik, RWTH Aachen, 52074 Aachen, Deutschland, Email: fme@akustik.rwth-aachen.de

Introduction

The noise generated by high speed moving vehicles is becoming the main annoyance to outside environments. Knowing the characteristics of sound sources and how the radiated noise sounds like will benefit traffic planning and noise control of ourdoor scenarios and urban environments. Auralization is a good way to achieve this aim. It is understood as the process of turning predictive or measured acoustic data, such as a noise spectrum, into an audible audio signal in Virtual Reality (VR) [1]. Auralization of the exterior sound of trains in virtual environments requires adequate arrangements of virtual sound sources and appropriate signals for those. Obtaining these is not a trivial task. A virtual train will only sound convincing if it is auralized with the necessary spatial extent and also the corresponding spatial arrangement of noise sources (e.g. wind noise, rails and wheels, pantograph). Beamforming has been successfully applied to identify and assess the relevant noise sources of trains [2]. Acoustic beamforming uses microphones to form an array to enhance the detected audio signal. The geometry and weighting of the array determine certain performance, resolution, signal-to-noise (SNR) ratio, directivity etc. When the received signal is spatially filtered, the signal from a particular steered angle is enhanced, which is usually used for the unknown source localization. Based on the localization of the sound sources and their characteristics (frequency components, amplitudes, phases etc.), can the synthesis be conducted for the auralization. For aircraft noise synthesis, a complete process is developed by Stephen A. Rizzi et al. [3], while for electrical railbound vehicles, a sound synthesis and validation model is proposed by M. Klemenz in [4].

This paper considers the application of beamforming for obtaining source signals for train auralization. Using a microphone array set up nearby a railway, pass-by noise of trains is recorded. In the subsequent post-processing, major sound sources are identified and the steered audios are generated using beamforming. Furthermore, the application of the sound source localization and steered audios tp train auralization is discussed.

Beamforming Theory

Figure 1 shows using a microphone array to measure a signal. There are two types of propagating waves, plane and spherical waves, considering the distance between source and receiver. $\mathbf{p}_n$ is the coordinates and $f(t, \mathbf{p}_n)$ is the received signal of the nth microphone. The time delay τ_n is expressed separately for plane and spherical waves:

$$\tau_n = \frac{\mathbf{a}^T \cdot \mathbf{p}_n}{c} \tag{1}$$

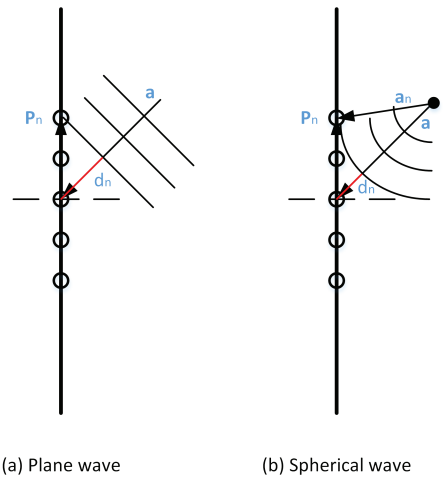

Figure 1: Traveling distance difference for plane and spherical waves.

$$\tau_n = \frac{\|\mathbf{a}_n\| - \|\mathbf{a}\|}{c} \tag{2}$$

where $d_n = \mathbf{a}^T \cdot \mathbf{p}_n$ for plane wave, and $\mathbf{a}$ is the unit vector of the wave propagating direction; while $d_n = \|\mathbf{a} + \mathbf{p}_n\| - \|\mathbf{a}\| = \|\mathbf{a}_n\| - \|\mathbf{a}\|$ for spherical wave, and $\mathbf{a}_n$ is the unit vector of wave propagating from the source to the nth microphone.

The delayed signals are summed up as

$$y(t) = \sum_{n=0}^{N-1} w_n f(t - \tau_n, \mathbf{p}_n) \tag{3}$$

where w_n is the weighting multiplied to the output signal from the nth microphone, N is the number of microphones. This algorithm is defined as delay and sum beamforming, or conventional beamforming [5].

Time delay τ_n corresponds to phase shift $\phi_n = e^{-j\omega\tau_n}$ [6] when the array output is transformed into frequency domain using Fourier Transform. For plane and spherical waves

$$\phi_n = e^{-j\mathbf{k}^T \mathbf{p}_n} \tag{4}$$

$$\phi_n = e^{-jk\|\mathbf{a}_n\|} e^{jk\|\mathbf{a}\|} \tag{5}$$

where $\mathbf{k} = \frac{\omega}{c}\mathbf{a}$ is the wave number, $k = \frac{\omega}{c}$ is the magnitude of $\mathbf{k}$.

Defining array manifold vector $\mathbf{v_k}$ as

$$\mathbf{v}(\mathbf{k}) = [e^{-j\mathbf{k}_T \mathbf{p}_0}; e^{-j\mathbf{k}_T \mathbf{p}_1}; ...; e^{-j\mathbf{k}_T \mathbf{p}_{N-1}}] \tag{6}$$

for plane wave, where $\mathbf{k}_T$ is the wave number with target direction;

$$\mathbf{v}(\mathbf{k}) = [e^{-jk\|\mathbf{a}_0\|} e^{jk\|\mathbf{a}_T\|}; e^{-jk\|\mathbf{a}_1\|} e^{jk\|\mathbf{a}_T\|}; ...; e^{-jk\|\mathbf{a}_{N-1}\|} e^{jk\|\mathbf{a}_T\|}] \tag{7}$$

for spherical wave, where $\mathbf{a}_T$ is the wave propagating unit vector in a target direction. Therefore, the array output expression in frequency domain is

$$Y(\omega, \mathbf{k}) = \mathbf{w}^T(\mathbf{k})\mathbf{F}(\omega) \tag{8}$$

where $\mathbf{w}_n(\mathbf{k}) = w_n \mathbf{v}_n(\mathbf{k})$, $\mathbf{F}_n(\omega)$ is the Fourier Transform of $f(t, \mathbf{p}_n)$. The above function is defined as frequency-wavenumber response function.

Measurement setup

A 24-microphone vertical array is set up 3.2 m from the near-side plane of the train. The microphones are put 0.08 m away from each other in a line. (See Figure 2 and Figure 3). Bombardier Talent 2 (Type 442) and Bombardier Talent (Type 644) are measured at the speed of 150 kph and 91 kph.

Figure 2: On-site measurement setup.

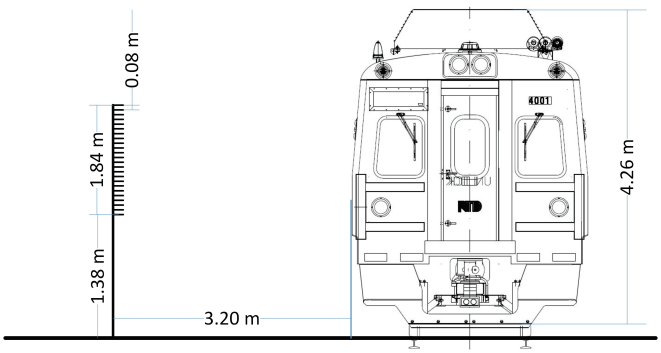

Figure 3: Front view of the measurement setup.

Source localization on the measured trains

Localization simulation on moving sound sources

In this work, uniform weighting is applied to the array for frequency lower than 2.5 kHz while Chebyshev weighting for frequency higher than 2.5 kHz. As is known, the array is expected to decrease the Maximum Sidelobe Level (MSL) and at the same time have narrow beamwidth, so to say to obtain better resolution. Applying Chebyshev weighting will reduce the MSL but meantime broaden beamwidth. However, although the beamwidth is increased, the resolution is still high enough to localize the sound sources in this research.

The microphone array has no horizontal resolution since it is vertical linear, however, considering the wave spherical makes sound sources still resolvable when they are just in front of the array. Thereby, the audio signal received by each microphone is divided into blocks. For each block, there is a part "facing" the array resulting in the reconstruction plane divided into corresponding parts (The principle can be seen in Figure 4). The reconstruction plane is regarded as the imaginary plane sticks to the near-side plane of the train. In each part, the array can be steered by applying different steering angles to search sources. In this way, each part is divided into grids for steering and so is the reconstruction plane.

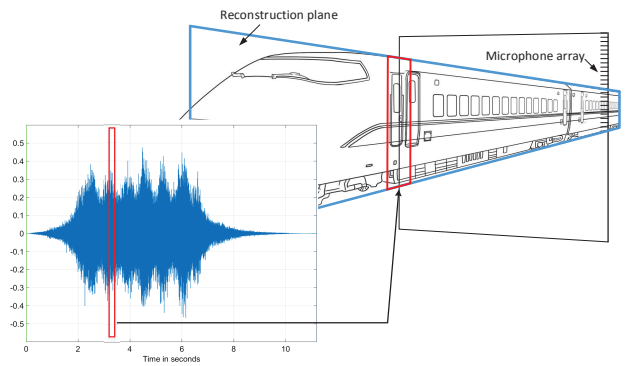

Figure 4: The principle to localize moving sound sources

A simulation is run using the Matlab ITA-Toolbox developed at the Institute of Technical Acoustics at RWTH Aachen University [7]. Three sources are located at a (1,3,0), b (0,3,5), c (-2,3,7) on a plane moving together towards negative z direction at speed of 1. Other setups are all the same as in the real measurement. (Figure 5) The moving plane is divided into uniform grids. Applying beamforming to each parts of the plane, the simulation result at 4 kHz is shown in Figure 6. As the array only has vertical resolution, the angle between the source and the array origin increases when a source is approaching and as the source leaving the angle decreases. Therefore it can be seen that there is a "v" shape around the source in the color map.

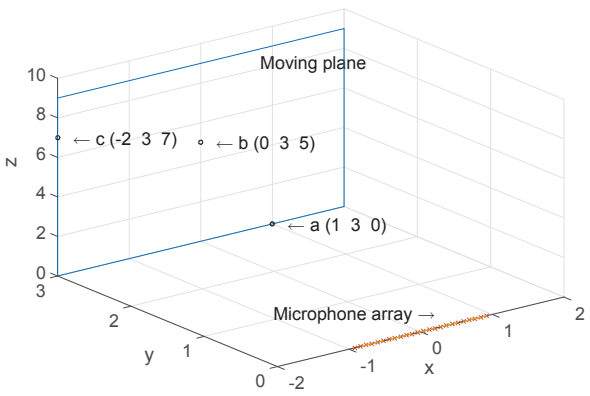

Figure 5: Microphone array and original positions of the moving sources in the simulation.

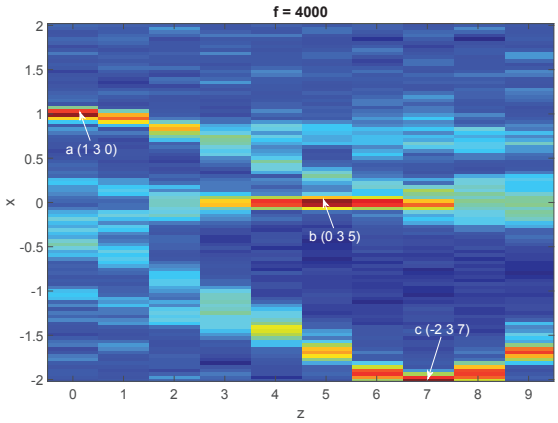

Figure 6: Simulation color map of three moving sound sources.

Localization results of the measured trains

As stated before, the reconstruction plane is divided into parts and beamforming is applied separately to localize moving sound sources. Hence the reconstruction plane of Bombardier Talent 2 is divided into 300 × 460 grids, and Bombardier Talent 118 × 260 grids as shown in Figure 7. The division is to make sure that there are 512 sample

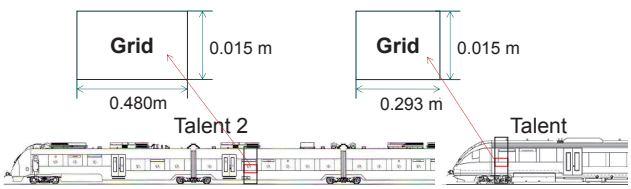

Figure 7: Grid division of the two trains.

in each block, guaranteeing the horizontal resolution and enough time information in each block for the frequency domain beamforming.

$$L_{grid} = \frac{v \times BlockSize}{fs} \tag{9}$$

where L_{grid} is the length of each grid, v is speed of train, fs is sampling rate (44.1 kHz). The color maps of the two measured trains are in Figure 8 and Figure 9. For Bombardier Talent 2, wheel is one of the main sound sources and the rail/wheel contact generates broadband noise ; as frequency increasess, pantograph, gaps between coaches and facilities on the top gradually become main sound sources; at around 1 kHz, the train head can be also regarded as main source. For Bombardier Talent, the main sound sourcesa are the same with the previous one except that the engine and cooling fan are also main sources instead of pantograph.

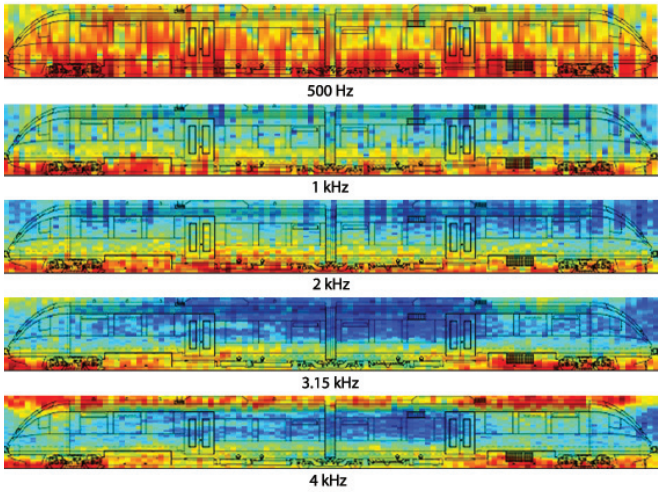

Figure 9: Source map of Bombardier Talent (Type 644).

Steered audio acquisition

Beamforming method is also used to created steered audios which play an important role as references in subsequent sound source synthesis. For instance, when a wheel as a sound source is to be synthesized, choose the audio file with the time slot when this wheel part is in front of the array and steer the angle to the wheel, therefore corresponding audio file is reconstructed after beamforming. The steered audio maintains more characteristics of a specific source and hence is a better reference for synthesis compared to the original recording. With source localization and synthesis, a virtual train can be created in virtual environment. Figure 10 shows placing sources on the virtual train. After obtaining source parameters and proper reproduction (using headphones or loudspeakers), auralization is then conducted.

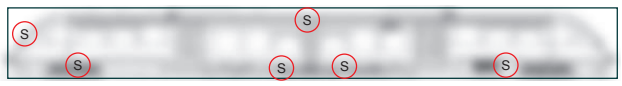

Figure 10: Sound sources on a virtual train.

Conclusion and outlook

A method based on delay and sum beamforming using vertical linear array to localize sound sources and obtain steered audios is developed and applied on two measured

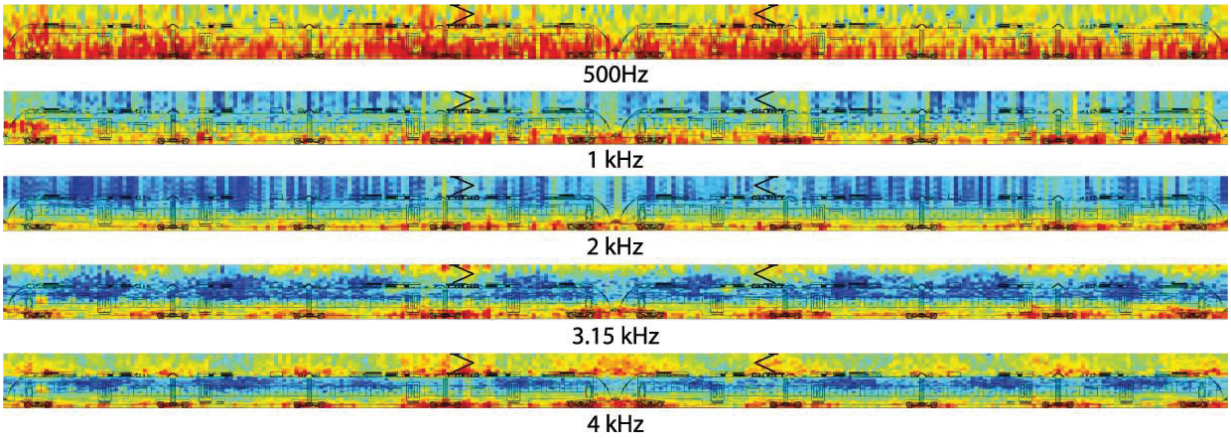

Figure 8: Source map of Bombardier Talent 2 (Type 442).

trains. Wheels are main sources radiating broadband noise; gaps between coaches, facilities on the top, pantograph (on Bambardier Talent 2) generate increasing aerodynamic noise as frequency increases; besides, engine and cooling fan are also main sources for Bombaidier Talent. Steered audios corresponding to different parts on the train are generated using beamforming as well to provide references for source synthesis. Further measurements on more trains, especially high-speed trains using larger and multidimensional array is needed to increase spatial resolution and array performance. Furthermore, the following synthesis will be done in tonal and broadband components separately with the knowledge of spectrums and steered audios. Finally, all the previous results will be applied in subsequent auralization.

References

[1] Sahai, A., Anton, E., Stumpf, E., Wefers, F., Vorlaender, M.: Interdisciplinary auralization of take-off and landing procedures for subjective assessment in virtual reality environments. 8th AIAA/CEAS aeroacoustics conference, 2012.

[2] Barsikow, B., King, W. F., Pfizenmaier, E.: Wheel/rail noise generated by a high-speed train investigated with a line array of microphones. Journal of Sound and Vibration 118.1 (1987): 99-122.

[3] Rizzi, S. A., Sullivan, B. M.: Synthesis of virtual environments for aircraft community noise impact studies. 11th AIAA/CEAS Aeroacoustics Conference. Vol. 4. 2005.

[4] Klemenz, M.: Sound synthesis of starting electric railbound vehicles and the influence of consonance on sound quality. Acta acustica united with acustica 91.4 (2005): 779-788.

[5] Van Trees, H. L.: Optimum Array Processing: Part IV of Detection, Estimation and Modulation Theory. John Wiley & Sons, New Jersey, 2002.

[6] Johnson, D. H., Dudgeon, D. E.: Array Signal Processing: Concepts and Techniques. Prentice Hall, New Jersey, 1993.

[7] Mueller-Trapet, M.: Comparison of Sound-Source Localization Methods for Vibrating Structures. RWTH Aachen University, 2009.