

MASG: MICROPHONE ARRAY SPEECH GENERATOR IN ROOM ACOUSTIC

Rui Cheng

Abstract

Microphone array speech generator (MASG) is used to simulate the speech data received by microphone array of various shapes in room acoustic environment, including clean speech (clean), reverberation speech (clean rever), noisy speech (clean rever noise), noisy and reverberation speech (clean rever noise) and corresponding noise signal (noise).

1 Method

The MASG is implemented based on two tools, namely, *Pyroomacoustic* [1] and an improved version *add_noise_for_multichannel* with add noise [2]. The schematic diagram of MASG is shown in Fig. 1.

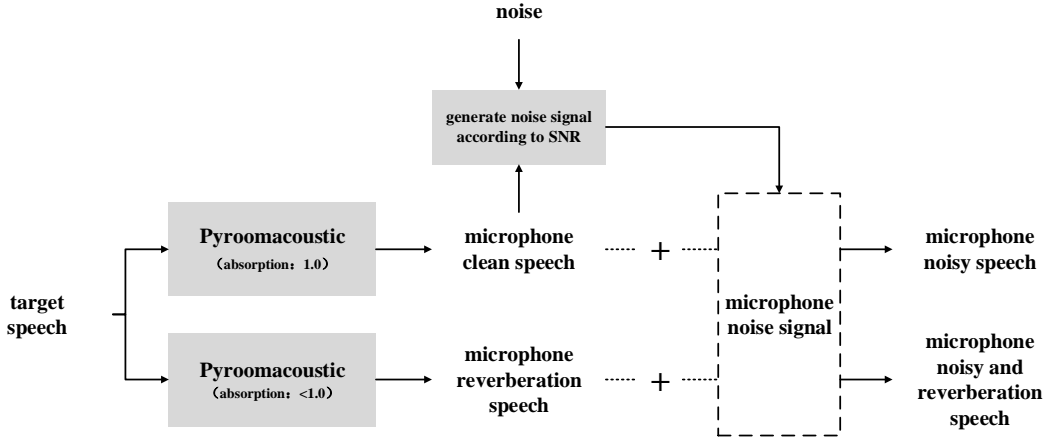


Fig. 1 The schematic diagram of MASG.

Based on *Pyroomacoustic*, the microphone array clean speech is obtained by setting the absorption to 1.0, and the microphone array reverberation speech is obtained by setting the absorption to less than 1.0. With the microphone array clean speech, combined with the noise signal and the expected signal-to-noise ratio (SNR), we can get the corresponding microphone array noise signal, and combine them with the microphone array clean speech and the microphone array reverberation speech to get the microphone array noisy speech and the microphone array noisy reverberation speech.

From this, we can get the simulation data of all microphone arrays used in indoor acoustic environment.

2 Simulation Environment

In order to verify the effect of the MASG, we set up a common room acoustic environment in our life, meeting room scene. This scenario is shown in Fig. 2.

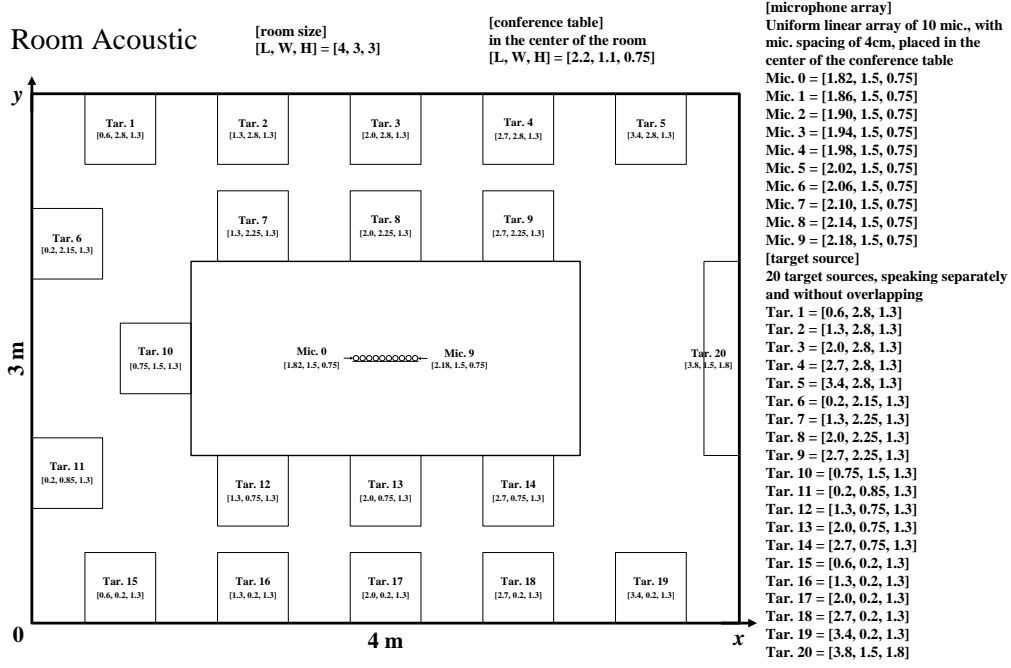


Fig. 2 Meeting room acoustic environment.

The scene simulates a meeting room with a length of 4m, a width of 3m and a height of 3m. In this room, a 2.2m×1.1m×0.75m conference table, 19 chairs with possible target sound source, and an audible screen are respectively placed. Their coordinates and details are shown in Fig. 2.

Based on such a meeting room environment, we abstract the room, microphone array, target source and other information used to make the data set, and get the simulation environment as shown in Fig. 3.

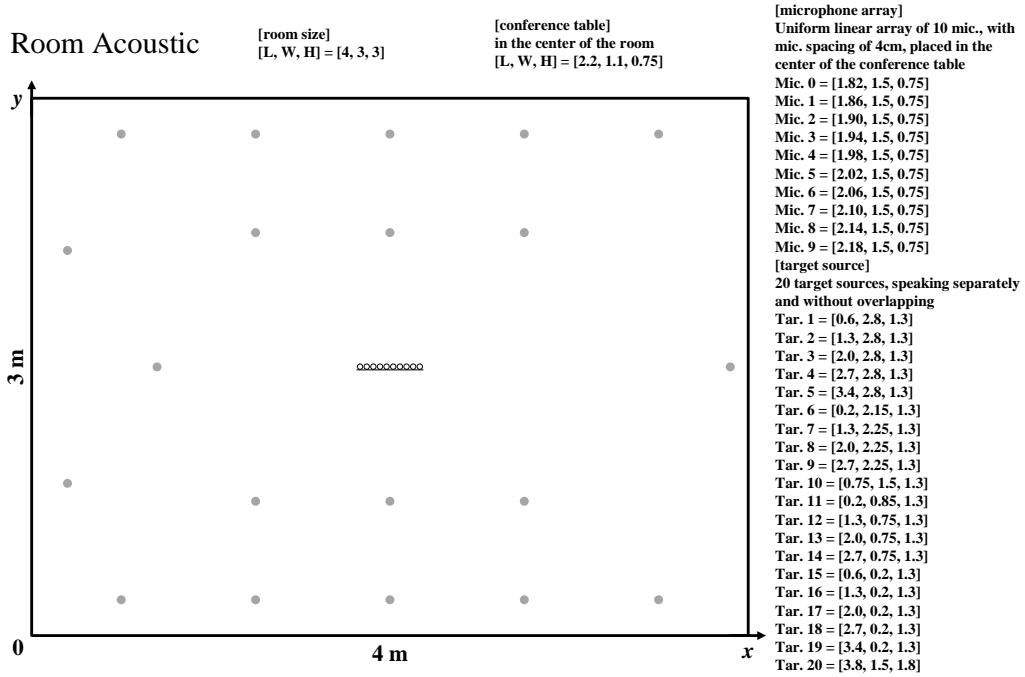


Fig.3 The simulation environment.

3 Program List

The MASG is implemented with Python. The detailed packages and functions are as follows.

3.1 Packages

[numpy]

<https://numpy.org/>

<https://pypi.org/project/numpy/>

[matplotlib]

<https://matplotlib.org/>

<https://pypi.org/project/matplotlib/>

[scipy]

<https://www.scipy.org/>

<https://pypi.org/project/scipy/>

[pyroomacoustic]

<https://github.com/LCAV/pyroomacoustics> <https://pypi.org/project/pyroomacoustics/>

3.2 Functions

[add_noise_for_multichannel.py]

This function is used to add noise to microphone array clean speech and microphone array reverberation speech based on the expected SNR.

[microphone_array_speech_generator_for_test_dataset.py]

This function is used to generate a microphone array speech test dataset for room acoustic environment.

[microphone_array_speech_generator_for_train_dataset.py]

This function is used to generate a microphone array speech training dataset for room acoustic environment.

[speech_connection.py]

This function is used to implement speech connection.

4 References

- [1] R. Scheibler, E. Bezzam and I. Dokmanić, "Pyroomacoustics: A Python Package for Audio Room Simulation and Array Processing Algorithms," *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, AB, 2018, pp. 351-355.
- [2] ITU-T (1993). *Objective measurement of active speech level*. ITU-T Recommendation P. 56.