Data clustering analysis of early reflections in small room
Zhaoqi Zhang, Ge Zhu, and Yong Shen

Citation: The Journal of the Acoustical Society of America 144, EL328 (2018); doi: 10.1121/1.5065073
View online: https://doi.org/10.1121/1.5065073
View Table of Contents: https://asa.scitation.org/toc/jas/144/4
Published by the Acoustical Society of America

ARTICLES YOU MAY BE INTERESTED IN

Composite honeycomb metasurface panel for broadband sound absorption
The Journal of the Acoustical Society of America 144, EL255 (2018); https://doi.org/10.1121/1.5055847

Highly directional pressure sensing using the phase gradient
The Journal of the Acoustical Society of America 144, EL346 (2018); https://doi.org/10.1121/1.5065401

Spectral directional cues captured by hearing device microphones in individual human ears
The Journal of the Acoustical Society of America 144, 2072 (2018); https://doi.org/10.1121/1.5056173

Integral identities for reflection, transmission, and scattering coefficients
The Journal of the Acoustical Society of America 144, 2109 (2018); https://doi.org/10.1121/1.5058681

Context integration deficit in tone perception in Cantonese speakers with congenital amusia
The Journal of the Acoustical Society of America 144, EL333 (2018); https://doi.org/10.1121/1.5063899

Green's function extraction by crosscorrelation and multidimensional deconvolution for outdoor sound propagation
The Journal of the Acoustical Society of America 144, EL353 (2018); https://doi.org/10.1121/1.5065072
Data clustering analysis of early reflections in small room

Zhaoqi Zhang, Ge Zhu, and Yong Shen
Institute of Acoustics, Nanjing University, Nanjing 210093, China
zhangzhaoqi03@gmail.com, gzhu06@outlook.com, yshen@nju.edu.cn

Abstract: It is common to increase the number of measurement points to improve the robustness of multipoint room equalization. However, the measurement of numerous points is extremely time-consuming and laborious. This letter analyzes the early reflections extracted from a large amount of room impulse responses using a $K$-means clustering algorithm, revealing that the spatial distribution of early reflections in the same cluster is not disorganized but regular and predictable. Furthermore, the results of the Monte Carlo simulation suggest that the appropriate selection of measurement positions can reduce the number of measurement points without compromising the robustness of multipoint room equalization.

1. Introduction

Room equalization is a crucial issue in various areas of applications, including home theaters, movie theaters, acoustic echo cancellation, and car HiFi systems. In the room equalization, room impulse response (RIR) characterizing the path from a sound source to a receiver is equalized with a suitably designed filter. However, the equalization filter designed to compensate for the room effects at a particular location performs poorly at other locations owing to significant variations of the RIRs in the receiver (listening) area. Hence, a multipoint room equalization method is proposed to improve the sound quality for all listeners. The multipoint equalizer is designed on basis of RIR measurements at few locations within the desired area, and a prototype response (PR) is typically obtained from these RIRs. The PR is representative of the listening area acoustic condition that requires correction. Certain PRs based on fuzzy C-means clustering are proposed in a previous work, where the number of measured RIRs is not sufficient to be persuasive, and the RIRs are divided into different groups without feature extraction. The previous effort inspires the authors to analyze a large number of RIRs with high spatial resolution using clustering techniques.

The robustness of multipoint equalization largely depends on the tolerance of the PRs to perturbations due to movement of the receivers, which can be described as the robustness of the multipoint measurement. It is common to increase the number of measurement points to acquire an accurate and robust PR. However, this process of measuring numerous RIRs within the listening area is time-consuming. A trade-off between reduction in the number of measurement points and improvement in the robustness of multipoint measurement has become an important issue, which needs to be investigated. This letter attempts to address this issue.

This letter measures a large number of RIRs in the listening area with high spatial resolution and analyzes the early reflections (ERs) using the data clustering method. The results reveal that the points with ERs of similar acoustic characteristics accumulate in space, which allows the utilization of one measurement point to represent a certain area. The comparison of two experimental results indicates that the spatial distribution of clusters changes according to the surrounding environment. Finally, a set of Monte Carlo simulations are performed to validate that the clustering results of ERs help simplify the measurement process without compromising the robustness of multipoint room equalization.

2. Analysis method

This section presents the methods applied in the RIR analysis. Section 2.1 describes the acoustic feature extraction techniques including RIR truncation and envelope calculation. Section 2.2 presents a general introduction of the $K$-means clustering.
2.3 cluster validity index adopted in this work to determine the optimal number of clusters. Finally, Sec. 2.4 outlines a Monte Carlo simulation utilized to analyze the robustness of multipoint measurement.

2.1 Feature extraction

To improve the performance of the data clustering algorithm, the position-related part needs to be separated from the RIRs and the acoustic feature must be extracted.

In room acoustics, an RIR can be divided into three parts: direct sound, early reflections, and late reflections. Direct sound is primarily determined by the distance between the sound source and the receiver, which carries little acoustic information of the listening area. The ERs consist of strong isolated reflections created by sound interacting with objects and surfaces in the room, which are strongly associated with the location of the measurement points and the surrounding acoustic environment. Late reflections are generally defined as a physical phenomenon, where a sufficient number of reflections overlap at any time along the time-axis; this part of the RIR can be modeled as a stochastic function with a normal distribution, which is position-independent. Therefore, the ERs are the major part of the RIR that varies significantly with respect to the position.

This letter extracts the ERs from an RIR based on the onset time, determined by a modified running local kurtosis analysis, and a duration of 60 ms. Then, the envelope of early reflections (EER) is calculated in the time domain through Hilbert transformation to remove the insignificant details. Let $h_{ER}(t)$ denote the ER part of the RIR; its envelope is expressed as

\[
E_{ER}(t) = \sqrt{(h_{ER}^2(t) + \hat{h}_{ER}^2(t))},
\]

where $\hat{h}_{ER}(t)$ represents the Hilbert transform of $h_{ER}(t)$, and is defined as

\[
\hat{h}_{ER}(t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{h_{ER}(t-t')}{t'} dt'.
\]

2.2 K-means clustering

After data feature extraction, a K-means clustering algorithm is used to analyze the group of EERs. This algorithm is one of the most widely used clustering techniques for shape-based time-series clustering. It aims to partition $N$ observations into $K$ clusters by minimizing the intrACLuster sum of squares defined as

\[
J = \sum_{j=1}^{K} \sum_{E_{ER,i} \in C_j} \|E_{ER,i} - m_j\|^2,
\]

where $E_{ER,i}$ ($i = 1, 2, \ldots, N$) is the $i$th measured EER, $C_j$ ($j = 1, 2, \ldots, K$) represents the $j$th cluster, $m_j$ is the centroid of $C_j$, and $N_j$ is the number of members in the $j$th cluster.

The clustering process involves (i) initializing $m_j$ by sampling among EERs with the K-means++ algorithm, which selects the initial centroids for K-means clustering based on a distance-weighted probability distribution; (ii) computing square distance $d_{ij}$ between $i$th measured EER and $j$th centroid,

\[
d_{ij} = \|E_{ER,i} - m_j\|^2,
\]

and assigning each EER to the closest cluster according to the distance; (iii) updating $m_j$ and returning to step (ii) until the cluster assignments no longer change. After clustering, the ERs in the same cluster share similar acoustic characteristics. In other words, their envelopes are similar to one another.

As the K-means clustering is a heuristic algorithm, there is no guarantee that it will culminate in the theoretical minimum of $J$, and the result may depend on the initialization. Therefore, this letter repeats this algorithm 200 times and saves the result with the minimum $J$.

2.3 Cluster validity index

Before analyzing the data, the optimal number of clusters $K^*$ needs to be determined by calculating the validity index $\kappa$. One example of the validity index is the Xie-Beni cluster validity index $\kappa_{XB}$ given as
It defines the intercluster separation as the minimum square distance between any two cluster centroids, and the intracluster compactness as the mean square distance between each data object and its cluster centroid. Thus, a smaller $\kappa_{XB}$ implies more compact and better clustering, and the optimal number of clusters $K^*$ can be found with the minimum $\kappa_{XB}$.

2.4 Monte Carlo simulation

A Monte Carlo simulation ordinarily uses repeated random sampling to obtain a statistically reliable result, and this letter utilizes 10,000 times of sampling to analyze the robustness of the multipoint measurement.

The process of the Monte Carlo simulation is as follows: (i) select a few RIRs from the measured data to simulate the measurement in the listening area (the method of selecting the points is further discussed in Sec. 3.2); (ii) use these RIRs to calculate the RMS spatial averaging frequency response as the PR$^2$ (the 1/12 octave-band smoothing of the magnitude of the frequency response is also used to provide a fine resolution at low frequencies and a lower resolution at high frequencies); (iii) compute the mean absolute error $e_{\log}$ (Ref. 12) (in dB) between the PR and the target response calculated with all data points; and (iv) repeat steps (i)–(iii) 10,000 times to simulate vast multipoint measurements and acquire a set of $e_{\log}$.

Multipoint measurement with strong robustness should have similar PR when varying the measurement positions. Specifically, the values of $e_{\log}$ tend to be close to one another. Hence, the standard deviations of the set of $e_{\log}$ can be used to evaluate the robustness of the multipoint measurement.

3. Experimental results and discussions

Figure 1(a) depicts the top view of the room, where two experiments are carried out in different acoustic environments. In experiment 1, only a sofa is placed in the listening area, whereas in experiment 2, a table and a small sofa (drawn in dashed lines) are set around the measurement region. The listening area covers not only the sitting area on the sofa but also the area in front of the sofa, which is also the activity area of the listeners. Besides, a large listening area with certain spatial sampling interval results in a large number of measurement points, which makes the results of the clustering more reliable and revealing. Each experiment measures 462 points with an interval of 0.12 m in the listening area using a microphone moved by a stepping-motor driving system, and in both cases, the locations of the measuring points remain the same.

3.1 Clustering results and analysis

This section presents the clustering results with comparison between the two experiments. The number of clusters is determined by the Xie–Beni index presented in Fig. 1(b), which shows that the optimal number of clusters is four in experiment 1 and five in experiment 2. In order to demonstrate the comparison of the results more clearly,
both experiments take four as the number of clusters. The choice is reasonable based on the fact that the value of $k_{XB}$ in experiment 2 remains small when the number of clusters is 4.

Figure 2 illustrates the results of clustering in three-dimensional space. The coordinate system is set according to Fig. 1(a) and the z-axis zero point is on the ground. As shown in Fig. 2, the ERs in the same cluster gather spatially rather than spread out in space. Therefore, it provides evidence that it is sufficient to measure the RIR at one point in a given area instead of choosing several points randomly in the area.

Comparisons of Fig. 1(a) with Figs. 2(a) and 2(b) reveal that the spatial distributions of the clusters correspond to the surrounding acoustic environment. For example, in the first experiment with a single sofa in the listening area, the acoustic environment remains the same along the y-axis. Consequently, the boundaries of the clusters are all parallel to the y-axis. In contrast, after adding the table and small sofa around the listening area, the cluster boundaries near them become tilted toward the y-axis, whereas the boundaries near the larger sofa remain parallel to the y-axis. In particular, the points near the table, which belong to different clusters in the first experiment, turn into the same cluster. In conclusion, the measurement locations can be selected according to the furniture setting to ensure coverage of larger number of clusters.

3.2 Results of Monte Carlo simulations and analysis

Based on the representative characteristics of clustering, a robust and accurate PR can be extracted from a few points on the condition that they are selected from different clusters. To verify the proposed viewpoint, Monte Carlo simulations are performed using the data of experiment 2 considering that the furniture setting in experiment 2 is more complex.

In this work, the $K$-means clustering algorithm divides the preprocessed data into five clusters, and then the set of $e_{log,1}$ are acquired through the Monte Carlo simulation outlined in Sec. 2.4 (the RIRs are sampled from each cluster randomly and every cluster provides one RIR). In contrast, the set of $e_{log,2}$ are obtained through another Monte Carlo simulation (the RIRs are selected from the entire listening area randomly, and the number of sampling points remains 5).

Figure 3(a) shows the statistical histograms of $e_{log,1}$ and $e_{log,2}$. The standard deviations of the sets of $e_{log,1}$ and $e_{log,2}$, denoted as $\sigma_1$ and $\sigma_2$, are calculated to be 0.1249 and 0.2176, respectively. Furthermore, the distribution of $e_{log,1}$ is more concentrated and right-skewed than that of $e_{log,2}$. It reveals that the PR calculated by RIRs selected from different clusters is more accurate and insensitive to varying measurement positions than that calculated by RIRs sampled randomly from the entire listening area. Namely, classifying the RIRs through the EERs is reasonable and effective, and the RIRs in the same cluster are equivalent in estimating the target room response.

In order to further verify the improvement in robustness of the multipoint measurement by selecting points in each cluster, the number of sampling points in the second Monte Carlo simulation is increased from 6 to 20. Each corresponding Monte Carlo simulation is also re-performed to acquire a set of $\sigma_{2,n}$ ($n = 6, 7, \ldots, 20$), where $\sigma_{2,n}$ represents the $\sigma_2$ recalculated with $n$ points.

![Fig. 2. Results of clustering in three-dimensional space. In each figure, the points in the same cluster are marked with the same color. Coordinate system is set according to Fig. 1(a) and the z-axis zero point is on the ground.](image)
Figure 3 depicts the comparison between $r_2$, $n$ and $r_1$ (using only five points). When the number of points reaches 16, $r_2$ is approximately equal to $r_1$, which implies their similar robustness. In conclusion, an appropriate selection of points distinctly helps acquire an accurate and robust PR with a small number of measurement points.

4. Conclusion

This letter measures 462 RIRs in the listening area with a high spatial resolution, and two scenarios with different furniture settings are investigated. The ER parts are analyzed using the $K$-means clustering algorithm as well as the Monte Carlo simulation. The clustering results illustrated in Fig. 2 exhibit that the ERs in the same cluster are spatially close to one another and their spatial distribution is associated with the surrounding furniture, which provides evidence that it is sufficient to measure only one point in a certain area. Then, the results of the Monte Carlo simulation shown in Fig. 3 demonstrate that the appropriate selection of measurement positions, which ensures coverage of a larger number of clusters, can reduce the number of measurement points without compromising the robustness of multipoint room equalization. Further investigations can determine if the precise theoretical acoustic model of ERs can be modified to identify a detailed relationship between the spatial distribution of clusters and the surrounding acoustic environment; thus, the receivers can be located in different clusters more accurately.

References and links