GEOMETRICAL ROOM GEOMETRY ESTIMATION FROM ROOM IMPULSE RESPONSES

Tilak Rajapaksha[*], *Xiaojun Qiu*[*], *Eva Cheng*[*] *and Ian Burnett*[†]

[*] RMIT University, Melbourne, Australia
E-mail: {s3469983@student.rmit.edu.au, xiaojun.qiu, eva.cheng}@rmit.edu.au
[†] University of Technology, Sydney, Australia
E-mail: ian.burnett@uts.edu.au

ABSTRACT

Room geometry estimation from corresponding Room Impulse Responses (RIRs) has attracted much attention in recent years, and a key challenge is to find the first order image source locations from the RIRs under different environments. Unlike the existing approaches which require *a priori* knowledge of the room or require some ideal conditions, this paper proposes an intuitive geometrical method based on the acoustical image source model. The proposed approach does not need any *a priori* knowledge of the room, only the RIRs from one arbitrary source location to five arbitrary receiving locations. The first order image sources of the walls in a room are identified first, then the room geometry is estimated based on the wall locations using a geometrical approach. Simulations with 2D and 3D convex polyhedral rooms demonstrate the feasibility and the precision of the proposed approach is discussed.

Index Terms—Room impulse response, image source, time of arrival, room geometry estimation

1. INTRODUCTION

Room geometry information has wide applications in acoustic modelling, source localisation and separation, forensic applications, spatial sound reproduction and architectural design [1-4]. Room geometry estimation from corresponding Room Impulse Responses (RIRs) has attracted significant attention in recent years because the estimation of room geometry from acoustic measurements provides an alternative unobtrusive and inexpensive approach to physical measurement-based methods [5-12].

In existing approaches for room geometry estimation, the image source method is often applied to obtain the RIRs due to its low computational complexity [1]. To estimate the room geometry with the image source method, the first order image source location of a reflecting wall is used to estimate the position of the reflective wall, from which the room geometry can be derived. However, a key challenge to the approach is how to determine the first order image source locations from the RIRs.

By using a single RIR and based on the assumption that the first and second order reflections can be identified in the RIR, Moore et al. [5] and Dokmanic et al. [6] proposed methods to estimate the geometry of rectangular and convex polygon shaped 2D rooms. However, the theoretically accurate Times Of Arrivals (TOAs) and adjacent second order reflections required in [6] are not always available.

Multiple RIRs have been used to localize a single wall in 2D rooms by mapping the microphone and sound source locations and corresponding TOA into elliptical constraints [7]. Canclini et al. [8] extended the method to localize a wall in 3D rooms; however, the challenge in extending to multi-wall scenarios is matching the TOAs to a corresponding wall and identifying the first order TOAs in RIRs.

Filos et al. [9] proposed to apply the Hough transform to match the reflections in the RIR to a corresponding wall. By locating a sound source very close to a wall to obtain the first order reflection next to the direct sound in RIRs and repeating the same procedure for all walls, room geometry can be estimated [10]. However, this approach needs to measure a set of RIRs equal to the number of walls. The method is further developed in [11], but still requires placing the sound source near to at least one wall to identify the first order reflection in the RIRs.

In 2013, Dokmanic et al. proposed using the Euclidean Distance Matrix based method to localize image sources and using a vector decomposition approach to identify the first order image sources [12]. Although higher order image source location vectors can always be decomposed into lower order image source location vectors, the method may classify the higher order reflections incorrectly if the first order reflections corresponding to the lower order reflections are missed while collecting the RIR peaks.

This paper proposes a two-step intuitive geometrical method to estimate room geometry based on RIRs from one arbitrary source location to five arbitrary receiving locations in a room and the relative locations of the microphones, no *a priori* knowledge of the room or assumptions about identifying first order image sources are made.

In the remainder of this paper, Section 2 introduces the proposed geometrical approach for estimating image source locations and room geometry. Section 3 presents the simulation results for a series of symmetric and asymmetric convex polyhedral room shapes and discusses the precision of the method. The paper is concluded in Section 4 and the future directions are outlined.

2. PROPOSED METHOD

A RIR represents the response of a room to an impulsive sound, which contains direct sound and reflections from the walls, ceiling and floor [12]. To perform room geometry estimation, the image source locations need to be determined first, and then the first order image sources need to be identified. In the proposed two-step approach, the image source locations are estimated by estimating the distances between five microphones (four in 2D rooms) and the image source locations using the RIRs, then geometrical constraints are applied to estimate the room geometry based on the estimated image source locations.

2.1. Real and image sound source location determination

In a convex 3D room with four microphones A, B, C and D located at (x_A, y_A, z_A), (x_B, y_B, z_B), (x_C, y_C, z_C) and (x_D, y_D, z_D), the distances from these microphones to an unknown real or image sound source at $P = (x, y, z)$ are related to their locations by

$$\begin{cases} (x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2 = a^2 \\ (x - x_B)^2 + (y - y_B)^2 + (z - z_B)^2 = b^2 \\ (x - x_C)^2 + (y - y_C)^2 + (z - z_C)^2 = c^2 \\ (x - x_D)^2 + (y - y_D)^2 + (z - z_D)^2 = d^2 \end{cases} \quad (1)$$

If the distances a, b, c and d and the microphone locations are known, the sound source location can be calculated by solving Eq. (1). Assuming that the number of impulses in RIRs from microphones A, B, C and D are N_A, N_B, N_C, and N_D respectively, then the number of candidate sound sources are $N_A \times N_B \times N_C \times N_D$. The distance from each candidate sound source to each of the microphones can be obtained from its RIR by estimating the time of arrival. For some candidate sound sources, if there is no solution for Eq. (1), these candidate sound sources are removed from the candidate list.

However, some sound sources in the candidate list may not correspond to actual real or image sound sources. An extra microphone is thus proposed in this approach to remove incorrect candidate sound sources.

Assume that the extra microphone E is located at (x_E, y_E, z_E), which can be used with any three of the existing four microphones to form a new equation similar to Eq. (1). By following the same procedure, another four sets of sound source candidate lists can be obtained. By comparing the 5 sets of candidate sound source locations, the actual real and image sound source locations can be determined according to the source locations common to all sets. To further improve the accuracy, more microphones can be used.

2.2. Room geometry estimation

Based on the real and image source locations determined in

Section 2.1, an intuitive geometrical method is proposed in this section to estimate the room geometry. First, the real sound source location (corresponding to the estimated sound sources) is identified by finding the nearest source location to all microphones. Assume that the number of image sound sources determined in last section is N, labeled as ISL_1, ISL_2, ISL_3, …, ISL_N according to the distance between the image sound source and the real sound source. Using the fact that the nearest image source ISL_1 is always one of the first order reflections for any room shape, the first reflecting wall can be determined and confirmed as the plane bisecting the real sound source location and this first order image source location (ISL_1). Using the same procedure, a number of candidate walls can be obtained.

To check whether a candidate wall is a real physical wall of the room or not, the reflective point location of the candidate wall to an arbitrary microphone and sound source location with respective to the already confirmed wall or walls are examined. The reflective point can be located by intersecting a line from the microphone to the image source and the candidate wall. If the reflective point location of the wall and the real sound source location are on different sides of any confirmed walls, this candidate wall cannot be a real wall and should be removed from the candidate wall list. By repeating the process, all real walls of the room can be determined. The method is summarized in Table 1.

Table 1. The proposed room geometry estimation method

Inputs: 5 RIRs from one sound source location to 5 microphone locations and the relative location of the microphones

1. Determine and list all actual real and image sound source locations according to the approach described in Section 2.1.

2. Identify the real sound source location (S) by finding the nearest source location to all microphones from the source location list.

3. Label all the image source location as ISL_1, ISL_2, ISL_3, …, ISL_N sequentially according to the distance between the real and image sound sources.

4. Confirm the first wall (Wall$_1$) as the bisecting plane between sound source (S) and ISL_1.

5. Estimation the locations of the rest of walls:

    For $n = 2$ to N

    - Estimate candidate wall$_n$ as the bisecting plane between S and ISL_n
    - Find the reflective point P_n by intersecting a line from a microphone to ISL_n and the candidate Wall$_n$

    Check whether P_n and S are on the same side of all the confirmed walls
    - If Yes, the candidate wall is confirmed as a real wall
    - If No, the candidate wall is not a real wall and ISL_n is not a first order image source.
    End

6. Solve all the confirmed walls for room vertices

Output: Vertice coordinates of the room

A simple example is given in Fig. 1, where the real sound source location S is identified first by finding the nearest source location to all microphones. Then, the first 3 image sources ISL_1, ISL_2, ISL_3 are located according to the distances between each of the image sound sources to the real sound source. ISL_1 is the nearest image source, so $Wall_1$ is first determined and confirmed as the plane bisecting the real sound source location and ISL_1. Using the same procedure, $Wall_2$ and $Wall_3$ are then also obtained. The reflective points P_2 and P_3 for $Wall_2$ and $Wall_3$ to an arbitrary microphone can be located as described in Table 1. Because P_2 and the real sound source location are on the same side of the confirmed $Wall_1$, $Wall_2$ is confirmed to be a real wall of the room. But P_3 and the real sound source location are not on the same side of the confirmed $Wall_1$ and $Wall_2$, so $Wall_3$ cannot be a real wall.
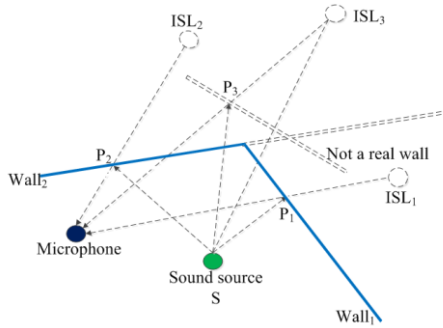

Fig. 1. Room geometry estimation with the proposed method

3. SIMULATIONS AND DISCUSSION

To evaluate the proposed room geometry estimation approach, a series of symmetrical and asymmetrical room shapes were investigated. The RIRs from one sound source to five microphones placed arbitrarily in the room were simulated using Odeon software. For the 2D rooms, four microphones were used, and the RIRs were simulated by setting the absorption coefficient of the floor and ceiling to full absorption. The accuracy of the estimation was evaluated by calculating the Euclidean distance between the estimated and real vertices of the rooms.

3.1. Image source location estimation

For the 2D triangular room shown in Fig. 2, the image source locations estimated with the proposed approach are presented in Fig. 3. The image source locations were estimated by using three sets of RIRs for microphone groups {A, B, C}, {A, B, D} and {A, C, D}. Fig. 3(a) shows the image source locations estimated by the {A,B,C} set, where there are many incorrect image source locations (blue stars) because only three microphones are used. With the extra microphone (D), Figs. 3(b) and 3(c) show the estimated image source locations from the {A,B,D} and {A,C,D} sets of microphones, respectively.
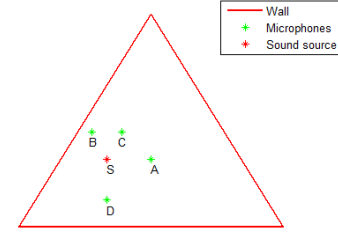

Fig. 2. 2D room geometry estimation for a triangular room with an arbitrarily located sound source and four microphones
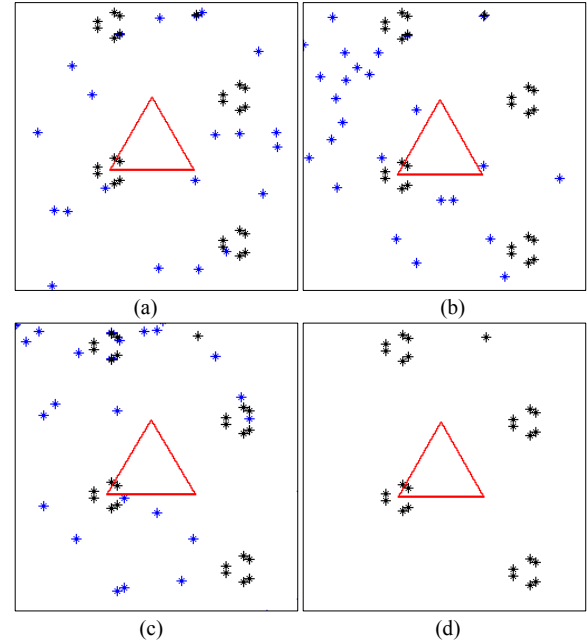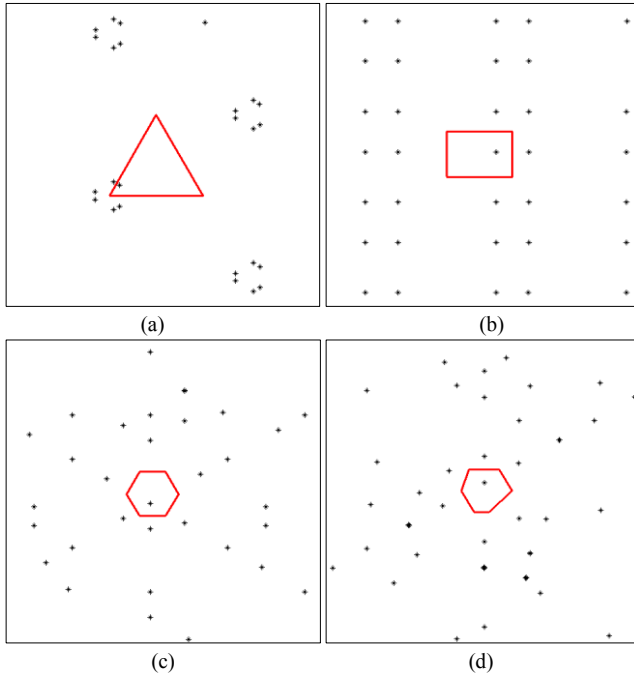

Fig. 3. Estimated possible image source locations for a 2D triangular room shown in Fig. 2 using microphone groups (a) {A, B, C} (b) {A, B, D} (c) {A, C, D}. (d) The image source locations estimated from locations common to (a), (b) and (c)

By considering all image source locations common to the three different sets of microphones, the actual sound image source locations can be estimated, as shown in Fig. 3(d). To further improve the estimation accuracy, the remaining combinations of microphones or more microphones can also be used. For estimating the image source locations in 3D rooms, at least five microphones are required.

3.2. Room geometry estimation

Fig. 4 illustrates four shapes of 2D rooms that were used to demonstrate the feasibility of the proposed method, where the estimated actual image source locations obtained with the proposed approach are indicated as black stars. Fig. 5 illustrates the four shapes of 3D rooms evaluated for demonstration. The corresponding room geometry estimation accuracy in terms of vertices estimation error obtained is shown in Figs. 6 and 7 for the corresponding 2D and 3D rooms, respectively.

Fig. 4. Simulation results for 2D rooms with different shapes (a) triangular; (b) rectangular; (c) hexagonal; (d) asymmetric hexagonal. The image source locations estimated according to Section 2.1 are also shown in the figure.
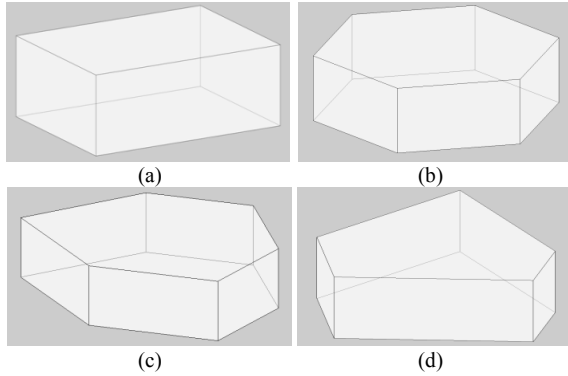

Fig. 5. Simulation results for 3D rooms with different shapes (a) rectangular shaped room; (b) hexagonal prism shaped room; (c) asymmetric hexagonal prism shaped room; (d) asymmetric pentagonal prism

Figs. 6 and 7 indicate that the maximum room geometry estimation errors are approximately 4.5 cm and 1 cm for the evaluated 2D and 3D rooms respectively. These indicative results were obtained from arbitrary sound source and microphone locations, where the room geometry estimation accuracy depends on the locations of the microphones and the sound source. In particular, if any three microphones are collinear in 2D rooms or any four microphones coplanar in 3D rooms, image sources cannot be estimated, the error is maximal and incorrect image source locations and room geometry are estimated.
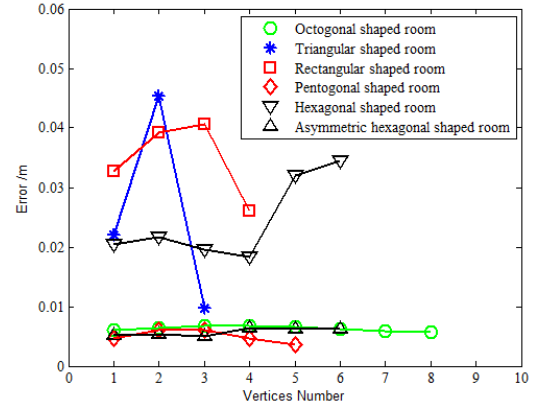
Fig. 6. Vertices estimation error for different 2D room shapes
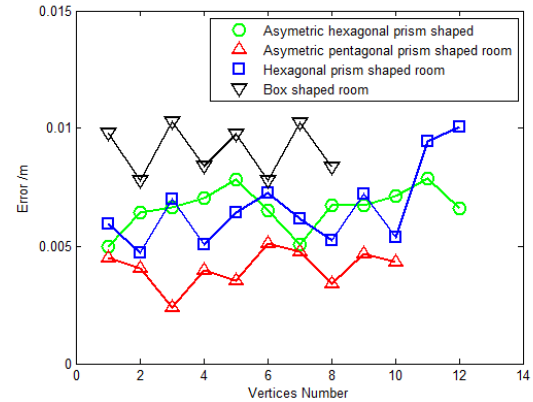
Fig. 7. Vertices estimation error for different 3D room shapes

4. CONCLUSION

This paper proposes an intuitive two-step geometrical method to estimate room geometry based on the acoustical image source model. Based on the RIRs from one arbitrary source location to five arbitrary receiving locations, the first order image sources of the walls in a room are identified first, then the room geometry is estimated based on the wall locations. Simulations with 2D and 3D convex polyhedral rooms demonstrate the feasibility of the method and the results indicate that the maximum room geometry estimation errors are approximately 1 cm for 3D rooms. Compared to the existing approaches, the proposed method does not depend on prior knowledge of the room (e.g., number of walls and room shape etc.), or the availability of higher order image sources, and does not restrict the angle between room walls. This paper addresses two key challenges in the current approaches, i.e., estimation of the sound source and all image source locations and identification of the first order image source locations. Future work includes investigating the relationship between room geometry estimation accuracy and sound source and microphone locations, and performing experimental verification.

REFERENCES

[1] J. B. Allen and D.A.Berkley, "Image method for efficiently simulating small-room acoustics," *Journal Acoust. Soc.America*, vol. 65, no. 4, pp. 943-950, Apr. 1979.

[2] F. Ribeiro, Z. Cha, D. A. Florencio, and D. E. Ba, "Using Reverberation to Improve Range and Elevation Discrimination for Small Array Sound Source Localization," *IEEE Transactions on, Audio, Speech, and Language Processing*, vol. 18, pp. 1781-1792, 2010.

[3] N. R. Shabtai, Y. Zigel, and B. Rafaely, "Feature selection for room volume identification from room impulse response," in *Proceeding of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, U.S.A., pp. 249-252, Oct. 2009

[4] T. D. Abhayapala and D. B. Ward, "Theory and design of high order sound field microphones using spherical microphone array," in *Proceeding of IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Orlando, Florida, U.S.A., pp.1949-1952, May 2002

[5] A. H. Moore, M. Brookes, and P. A. Naylor, "Room geometry estimation from a single channel acoustic impulse response," in *Proceeding of the 21st European Signal Processing Conference (EUSIPCO)*, Marrakech, Morocco, pp. 1-5, Sep. 2013.

[6] I. Dokmanic, Y. M. Lu, and M. Vetterli, "Can one hear the shape of a room: The 2-D polygonal case," in *Proceeding of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, pp. 321-324, May 2011.

[7] A. Canclini, E. Antonacci, M. R. P. Thomas, J. Filos, A. Sarti, P. A. Naylor and S.Tubaro, "Exact localization of acoustic reflectors from quadratic constraints," in *Proceeding of IEEE Workshop Applicat. Signal Process. Audio Acoust.(WASPAA),* New Paltz, New York, U.S.A., pp. 17-20, Oct. 2011.

[8] A. Canclini, F. Antonacci, J. Filos, A. Sarti, and P. Naylor, "Exact Localization of Planar Acoustic Reflectors in Three-Dimensional Geometries,"in *Proceeding of IWAENC International Workshop, Acoustic Signal Enhancement*, Aachen, Germany, pp. 1-4, Sep. 2012.

[9] J. Filos, A Canclini, M.R.P. Thomas, F. Antonacci, A. Sarti, and P.A. Naylor, "Robust inference of room geometry from acoustic measuremnts using the Hough transform," in *Proceeding of 19th European Signal Processing Conference (EUSIPCO)*, Barcelona, Spain, pp. 161-165, Aug.- Sep. 2011

[10] J. Filos, A. Canclini, F. Antonacci, A. Sarti, and P. A. Naylor, "Localization of planar acoustic reflectors from the combination of linear estimates," in *Proceeding of 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania, pp. 1019-1023, Aug.2012.

[11] F. Antonacci, J. Filos, M. R. P. Thomas, E. A. P. Habets, A. Sarti, P. A. Naylor and S. Tubaro, "Inference of Room Geometry From Acoustic Impulse Responses," *IEEE Transactions on, Audio, Speech, and Language Processing,* vol. 20, pp. 2683-2695, 2012.

[12] I. Dokmanic, R. Parhizkar, A. Walther, Y. M. Lu, and M. Vetterli, "Acoustic echoes reveal room shape," *Proceeding of the National Academy of Sciences of the United States of America*, vol. 110, pp. 12186-91, 2013.