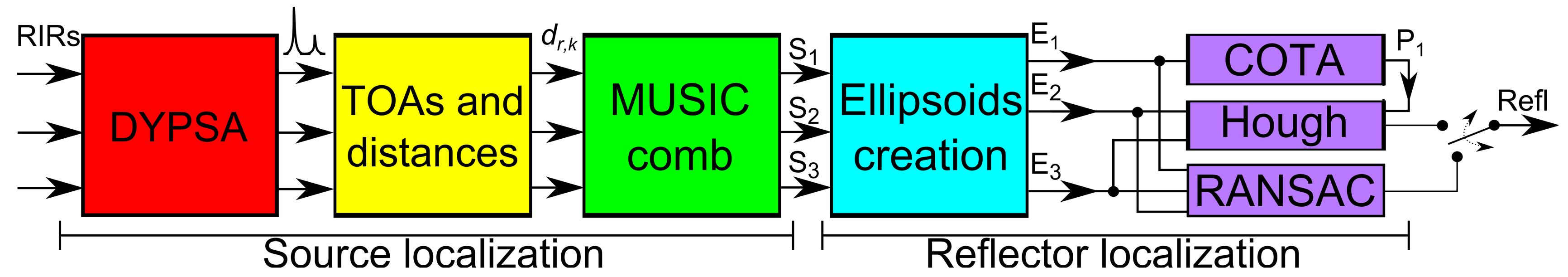
# A 3D Model for Room Boundary Estimation

Luca Remaggi, Philip J. B. Jackson, Wenwu Wang and Jonathon A. Chambers
Centre for Vision Speech and Signal Processing, University of Surrey, Guildford, UK

## Abstract

**Estimating the geometric properties of an indoor environment through acoustic room impulse responses (RIRs) is useful in various applications, e.g., source separation, simultaneous localization, and spatial audio. In this article, components of a method for environments reconstruction, that we have previously developed, have been extended for 3D: multiple signal classification (MUSIC) for direction of arrival (DOA) estimation, numerical search for reflector estimation and the Hough transform to refine the results. A variation is also proposed using random sample consensus (RANSAC) instead of the numerical search and the Hough transform which significantly reduces the run time. Both methods have been tested on simulated and measured RIRs. The proposed methods perform better than the baseline.**

## Model overview

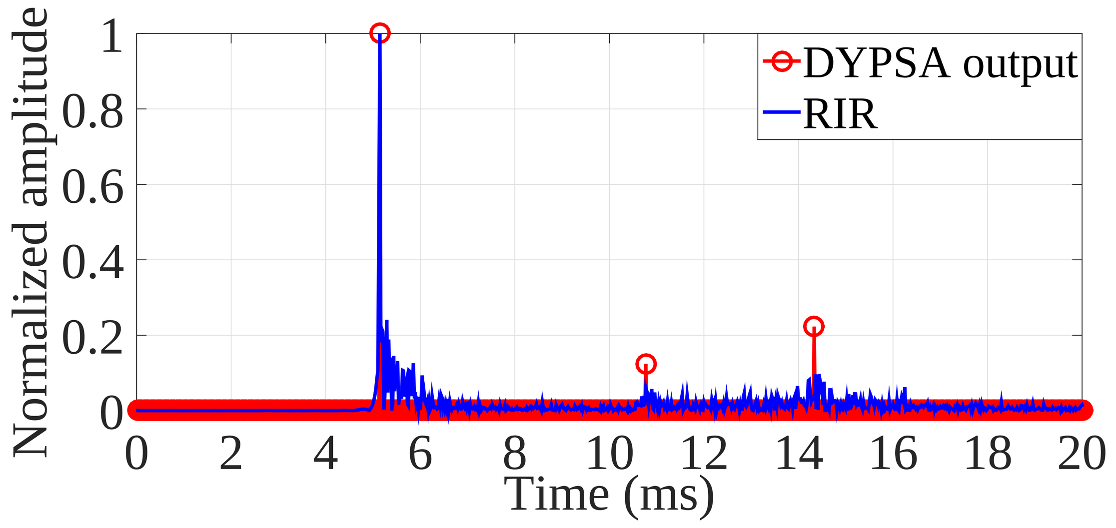



- In [1] we modify the algorithm in [2] introducing dynamic phase slope algorithm (DYPSA) for epoch detection, 2D MUSIC for DOA estimation and 2D common tangent algorithm (COTA) for reflector search
- Here, the novelty is given by the conversion of MUSIC, COTA and Hough in 3D, the production of ellipsoids instead of ellipses and the introduction of RANSAC as variant of COTA-Hough combination

## Source localization

### Distance calculation

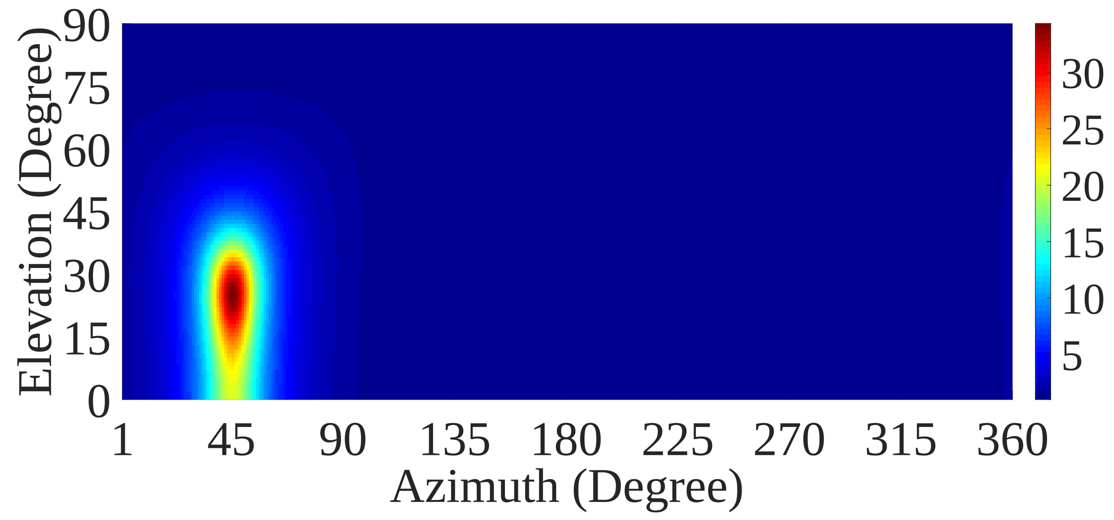
- RIRs peaks are detected using DYPSA




- Knowing $k$ time of arrivals (TOAs), the length of the sound paths considering the $r$-th mic are calculated $d_{r,k} = \tau_k \cdot c$
- The first peak $k = 0$ gives the distance between the source $u$ and the sensors array

### 3D MUSIC combination

- 3D MUSIC is exploited to estimate the 3D DOA of the signal respect to the uniform rectangular array (URA)

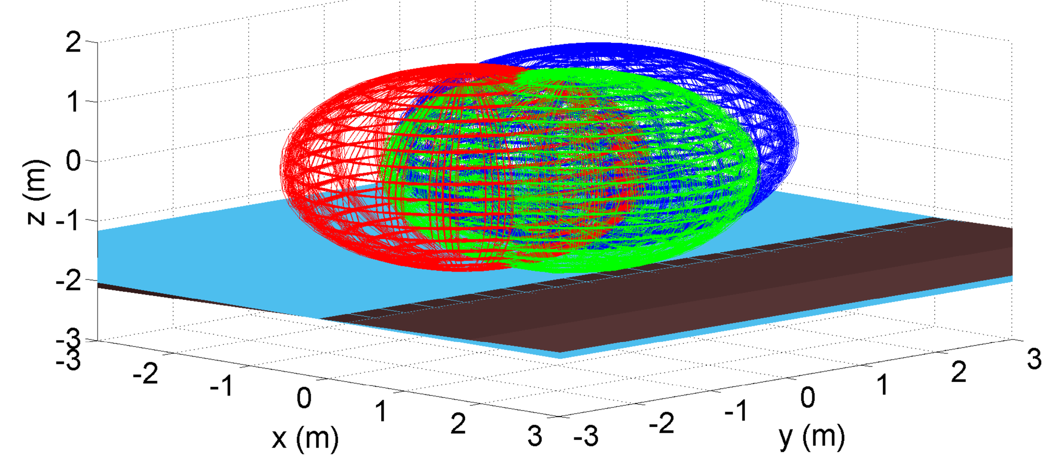



- DOAs (elevation $\Theta$ and azimuth $\Phi$) from MUSIC are combined with $d_{r,0}$ from DYPSA to implement a triangulation technique

## Reflector localization

### Ellipsoid generation

- Each ellipsoid has foci on the source $u$ and microphone $r$ and is defined by the matrix $\mathbf{E}_{r,u}$




- The major axis corresponding to the $k$-th reflection path length is $M_{r,k} = d_{r,k}$. The minor axes are equals and given by $m_{r,k} = \sqrt{M_{r,k}^2 - d_{r,0}^2}$

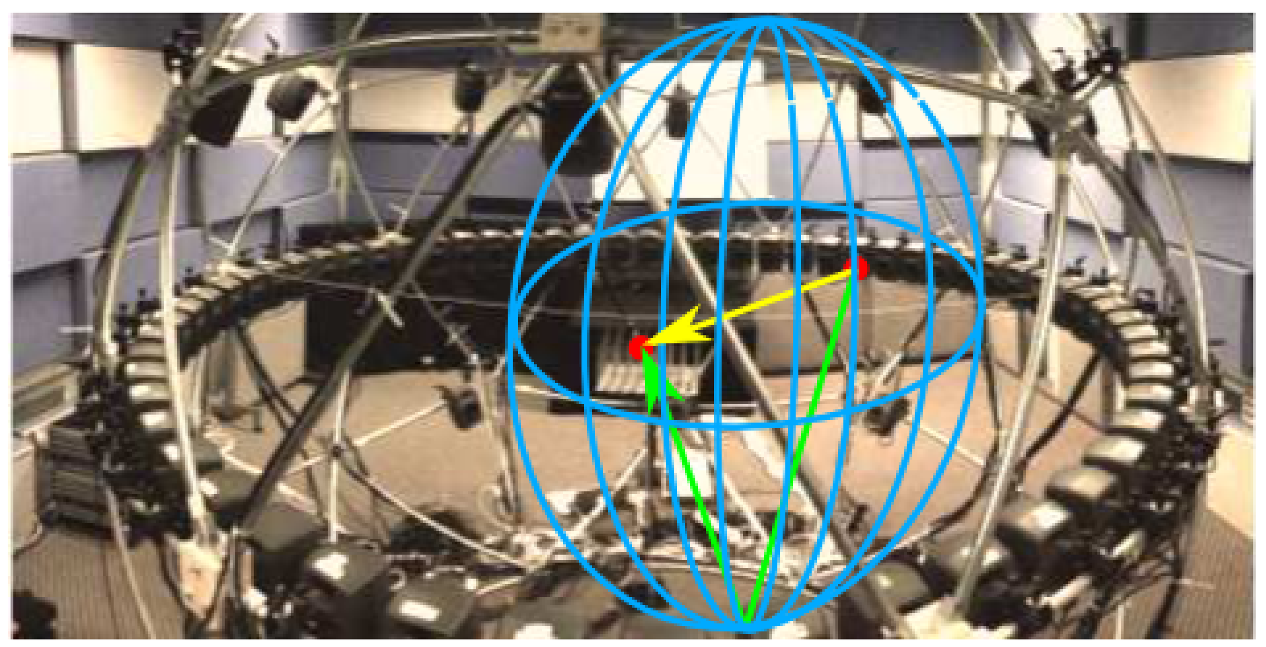
### 3D COTA - 3D Hough

- The plane $\mathbf{p}$ is the one common tangent to every ellipsoid $\mathbf{E}_{r,u}$
- The 3D-COTA finds $\mathbf{p}$ which minimizes $J(\mathbf{P}) = \sum_{r=1}^{M} |\mathbf{p} \cdot \mathbf{E}_{r,u} \cdot \mathbf{p}^T|^2$
- The Hough transform refines the result

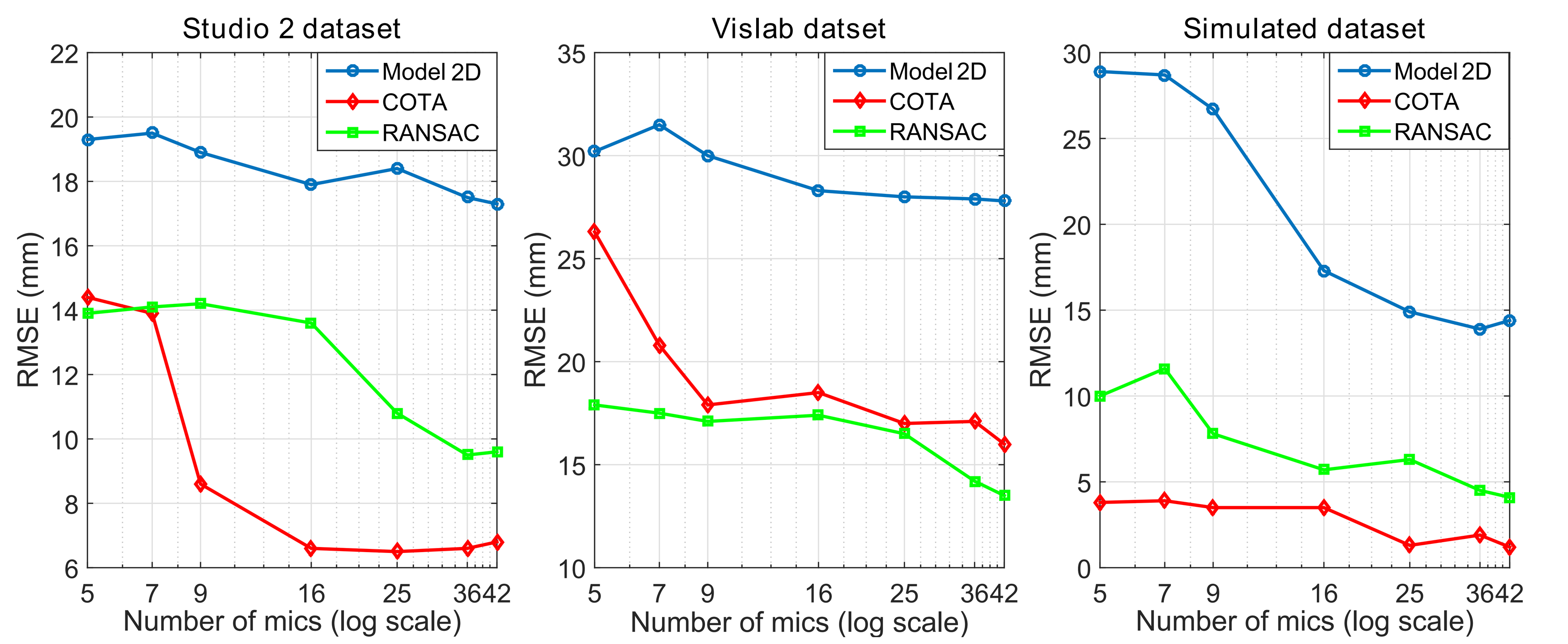
### RANSAC

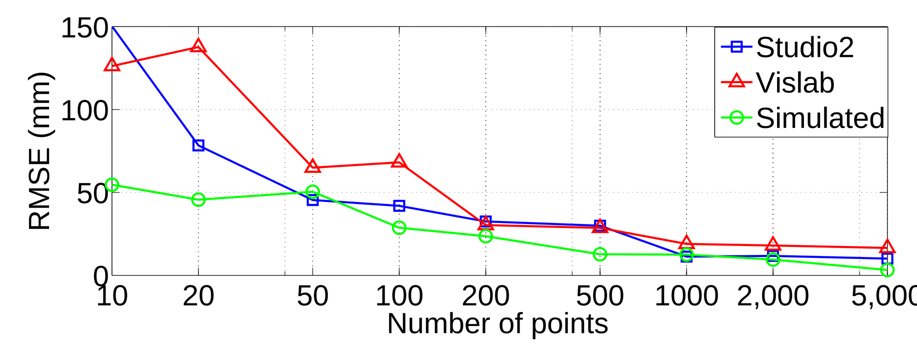
- A point lying on $\mathbf{E}_{\mathbf{r},\mathbf{u}}$ is randomly selected.
- To verify if the plane tangent in that point is tangent to all the other ellipsoids $|\mathbf{p}_l^T \mathbf{E}^* \mathbf{p}_l| = t_l$ is compared to a threshold.
- The plane that has the fewest non-tangent ellipsoids is selected as the estimated reflector.

## Experimental evaluation



- Studio-2 dimensions: $6.55 \times 8.78 \times 4.02$ m$^3$; RT60 235 ms
- Vislab dimensions: $7.90 \times 6.00 \times 3.98$ m$^3$; RT60 215 ms
- Simulated room dimensions: $14.55 \times 17.08 \times 6.05$ m$^3$; RT60 820 ms




(Top) RMSEs using different number of points (RANSAC) and three datasets
(Left) RMSEs for different number of microphones.

## Conclusion

Two versions of a model to estimate the room geometry have been presented. Tests for both simulated and real RIRs have been performed, and the results compared. The new 3D model showed a significant improvement over the 2D one. The use of RANSAC instead of COTA and Hough reduces 148 times the run time of the algorithm, thus it is recommended for fast processing. Future work will investigate datasets with different heights of microphones and sources, and extend analysis beyond the first reflection.

## References

[1] Remaggi et al., "Room boundary estimation from acoustic room impulse responses", SSPD, 2014.
[2] Antonacci et al., "Inference of room geometry from acoustic impulse response", IEEE Trans., 2012.

**Acknowledgments**:This work was supported in the UK by the Engineering and Physical Sciences Research Council (EPSRC) Grants EP/K014307/1 and EP/L000539/1, and the MOD University Defence Research Collaboration in Signal Processing.