Acoustic Reflector Localization: Image Source Reversion and Direct Localization

L. Remaggi, P. J. B. Jackson, P. Coleman and W. Wang

Centre for Vision, Speech and Signal Processing (CVSSP)

University of Surrey

Motivation

- Room geometry estimation provides several opportunities of improving different audio signal processing research areas.

- Estimating the room shape can improve areas exploiting environmental geometry such as:
  - Spatial audio[1,2]

- It can also help in enhancement of target signals:
  - Source separation[3]

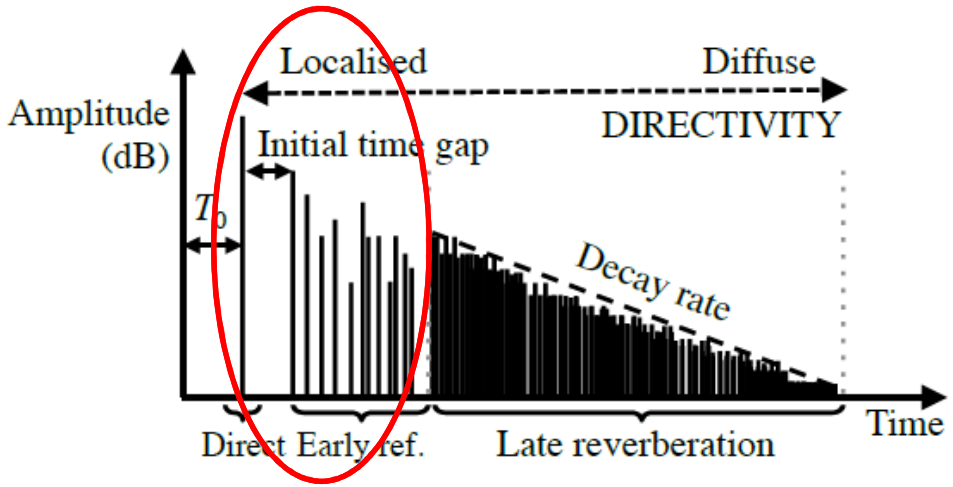
- In addition, it can be exploited to create robust hybrid models, together with image processing techniques[4]

UNIVERSITY OF SURREY

[1]Remaggi et al., AES Convention, 2015
[2]Coleman et al., JAES, 2017
[3]Asaei et al., IEEE TASLP, 2014
[4]Hussein et al., AAAI Conference, 2014

Room Impulse Response

- It can help with the acoustical behavior of rooms, extracting parameters. In fact, RIRs are composed by 3 components:
  - Direct sound, revealing the position of the source;
  - Early reflections, conveying a sense of the geometry;
  - Late reverberation, indicating the size of the environment. It does not have directional information.



- Direct sound and early reflection are so considered as providers of room geometry experimental setup parameters

UNIVERSITY OF SURREY

Image from [Coleman et al., AES Convention, 2016]

State-of-the-art classification

	Direct localization	**Image source reversion**
Models	• Kuster et al., 2004 • Nastasia et al., 2011 • Zamaninezhad et al., 2014 • Our ellipsoid tangent sample consensus (ETSAC)	• Tervo et al., 2012 • Ribeiro et al., 2012. • Dokmanić et al., 2013. • Our image source direction and ranging with loudspeaker image bisection (ISDAR-LIB) and variants

State-of-the-art classification

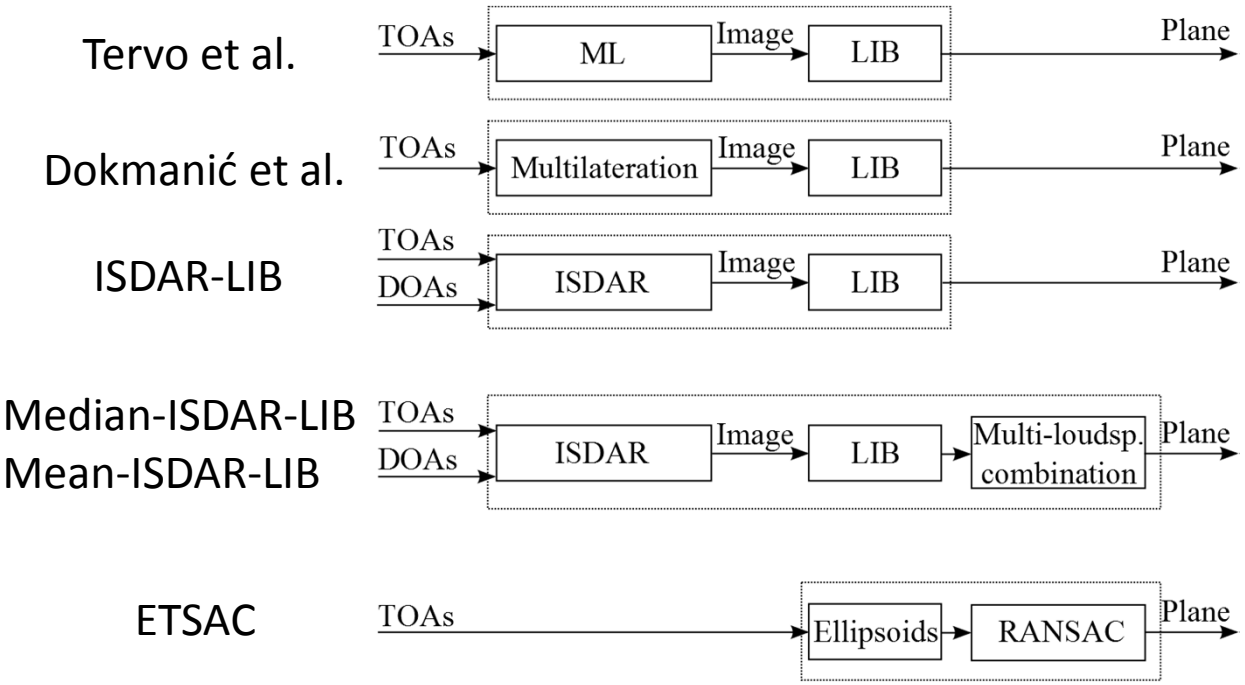
	Direct localization	Image source reversion
Models	• Kuster et al., 2004 • Nastasia et al., 2011 • Zamaninezhad et al., 2014 • Our ellipsoid tangent sample consensus (ETSAC)	• Tervo et al., 2012 • Ribeiro et al., 2012. • Dokmanić et al., 2013. • Our image source direction and ranging with loudspeaker image bisection (ISDAR-LIB) and variants

- Kuster[1] implied ULAs parallel to reflector
- Zamaninezhad[2] could localize two opposite reflectors only
- Ribeiro[3] required too huge datasets
- A comparison with Nastasia[4] is already state-of-the-art[5]

[1]Kuster et al., JASA, 2004
[2]Zamaninezhad et al., ISCCSP, 2014
[3]Ribeiro et al., IEEE TASLP, 2012
[4]Nastasia et al., EUSIPCO, 2011
[5]Remaggi et al., ICASSP, 2015

UNIVERSITY OF SURREY

The evaluated methods



The proposed methods

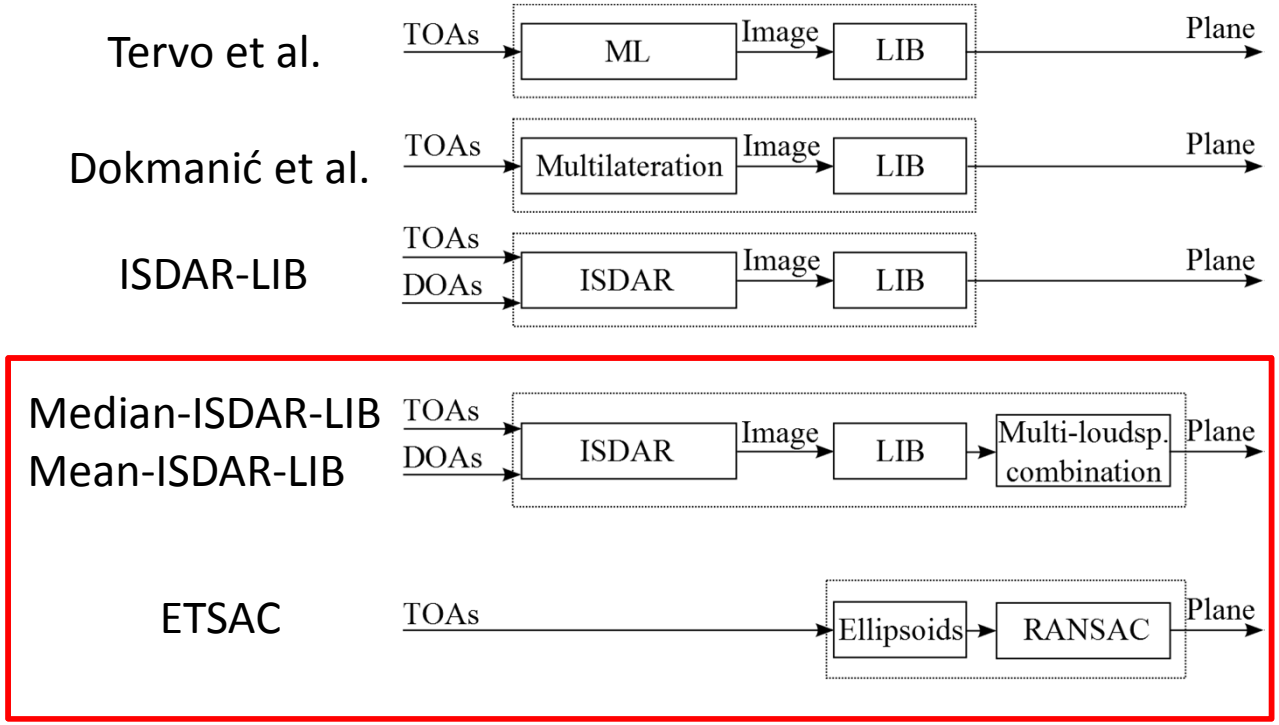


Image Source Reversion

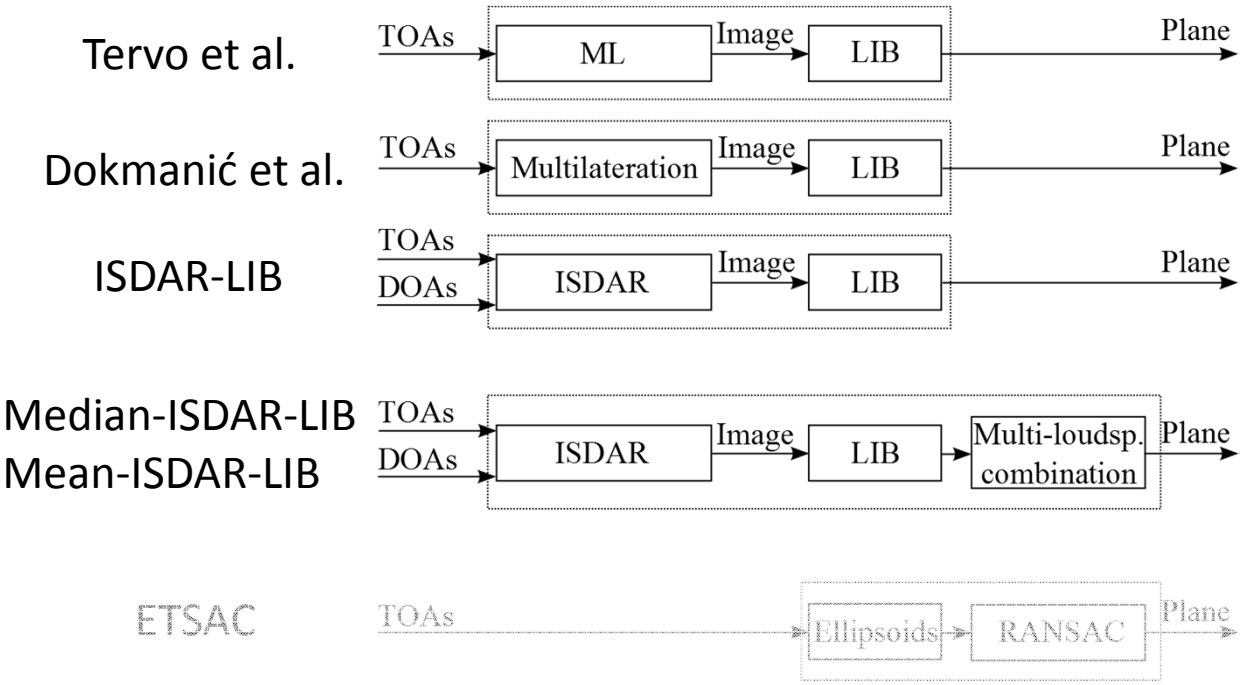
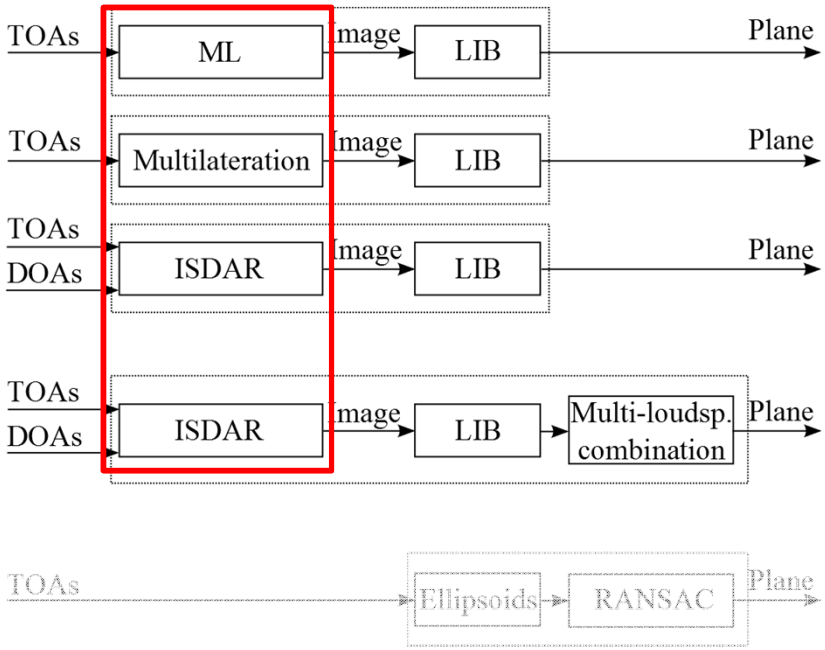


Image Source Reversion

- Image localization

- Maximum Likelihood[1]
- Multilateration[2]
- Image source direction and ranging (ISDAR)[3]



[1]Tervo et al., ICASSP, 2012
[2]Dokmanić et al., PNAS, 2013
[3]Remaggi et al., IEEE TASLP, 2017

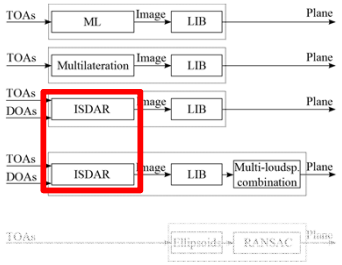
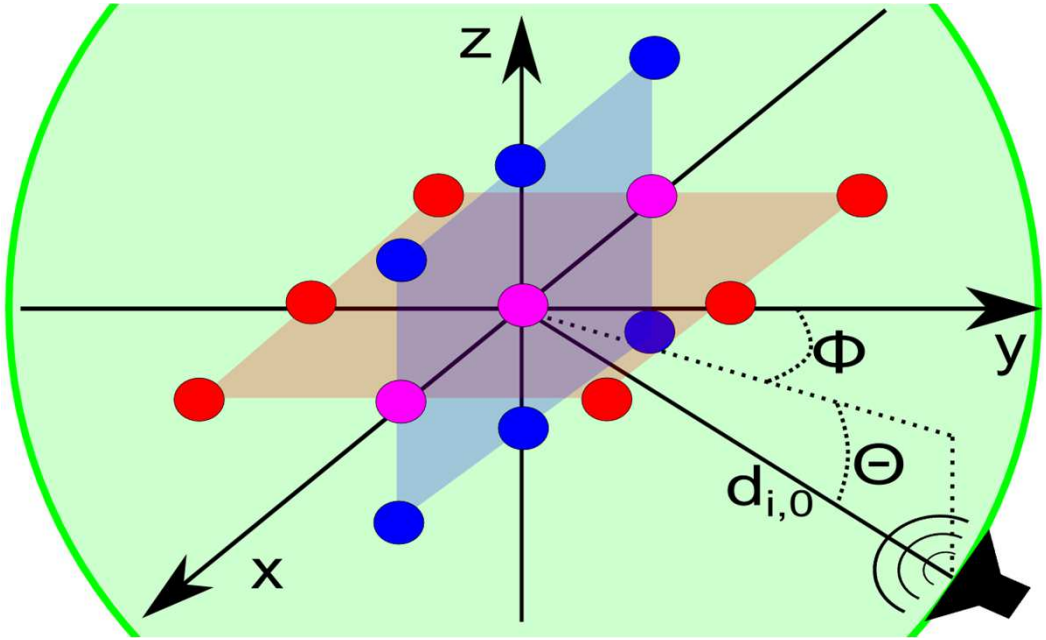


Image Source Reversion

- Image localization
  - ISDAR[1]



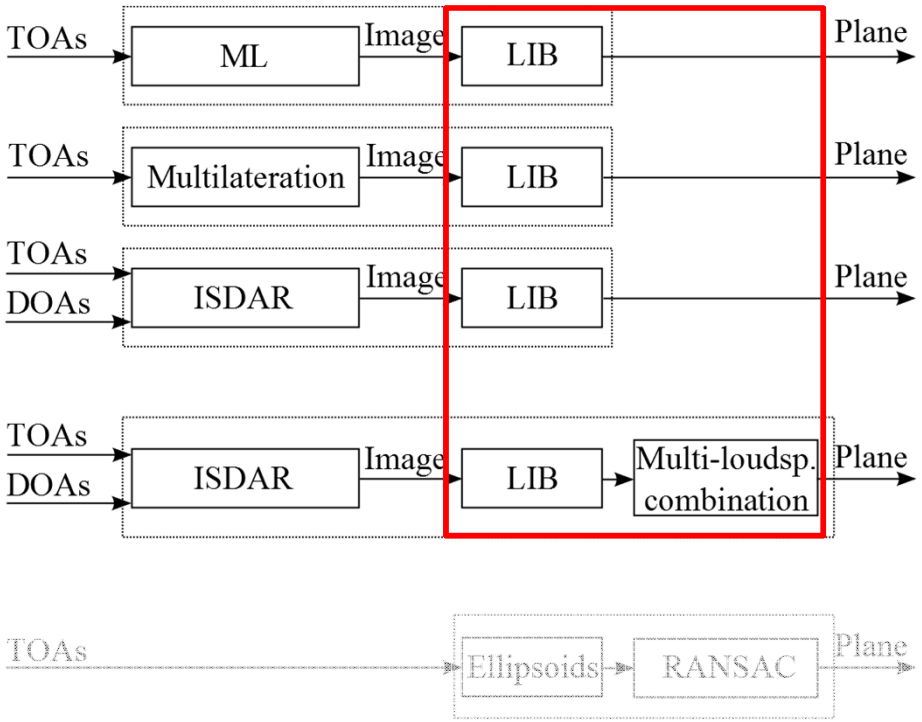
UNIVERSITY OF SURREY

[1]Remaggi et al., IEEE TASLP, 2017

Image Source Reversion

- Reflector localization

- Loudspeaker image bisection (LIB)[1]
- Loudspeaker image bisection (LIB)[2]
- Loudspeaker image bisection (LIB)[3]
- LIB + multi-loudspeaker combination[3]



[1]Tervo et al., ICASSP, 2012
[2]Dokmanić et al., PNAS, 2013
[3]Remaggi et al., IEEE TASLP, 2017

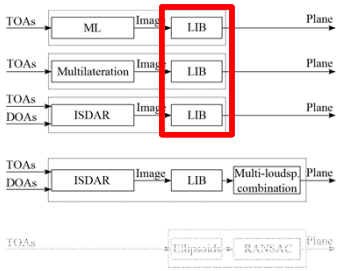
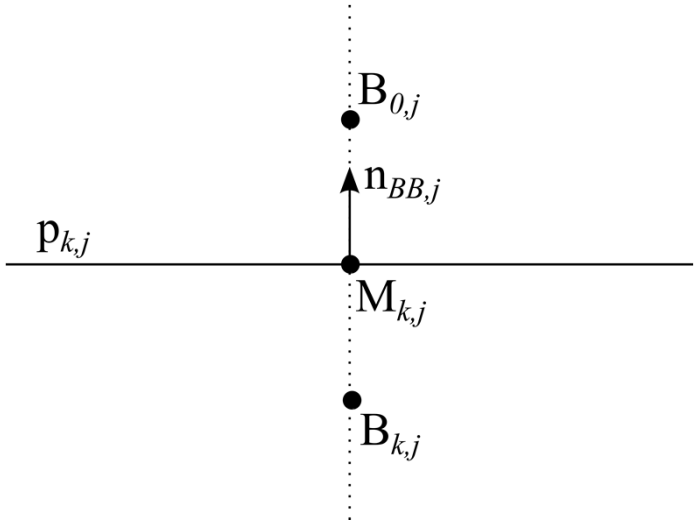


Image Source Reversion

- Reflector localization
  - LIB



  - ISDAR-LIB

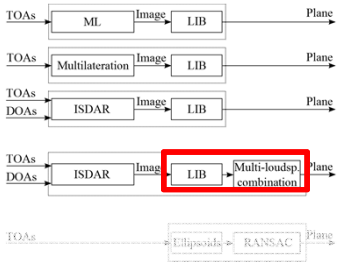
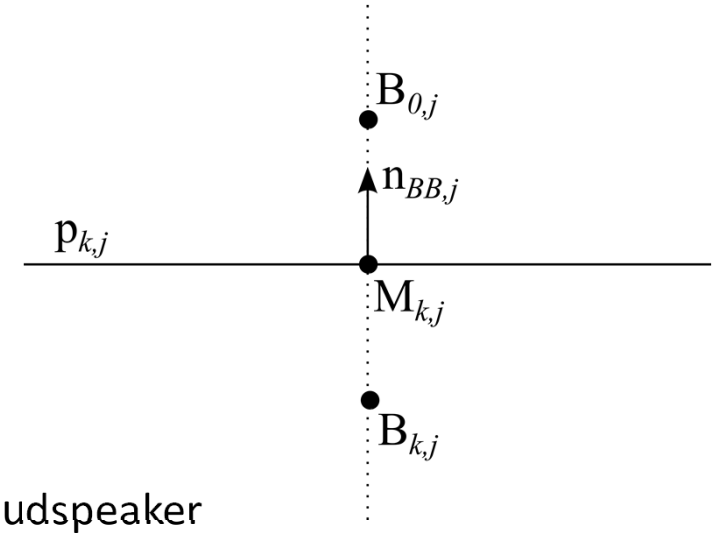


Image Source Reversion

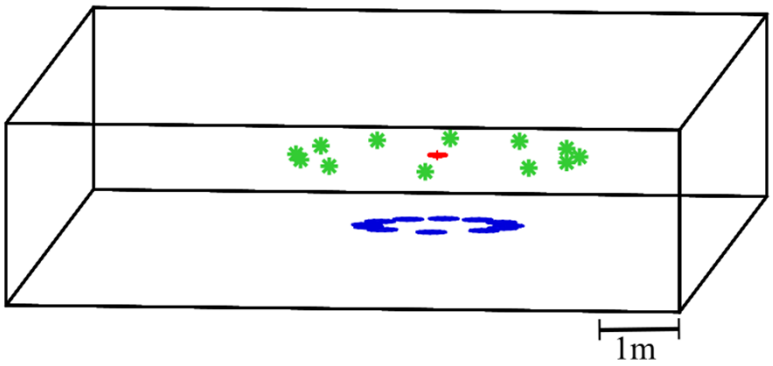
➢ Reflector localization

- LIB



- ISDAR-LIB

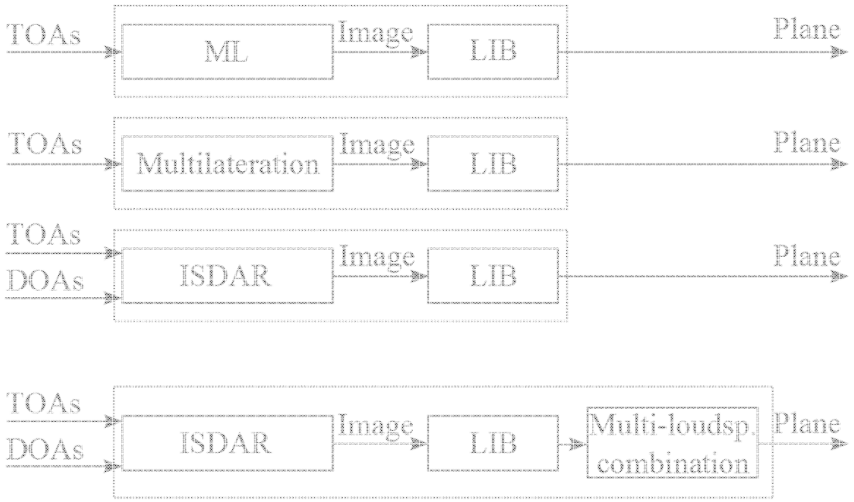
- LIB + multiple loudspeaker



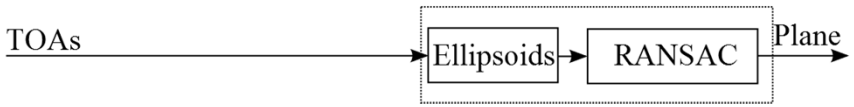
- Median-ISDAR-LIB
- Mean-ISDAR-LIB

UNIVERSITY OF SURREY

Direct Localization



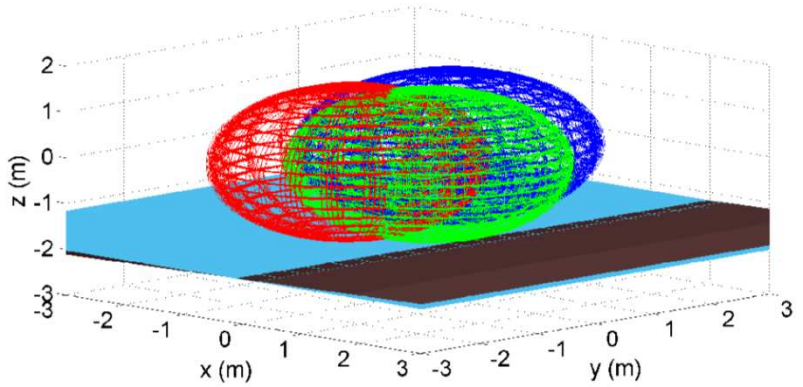
- Ellipsoid tangent sample consensus(ETSAC)[1]

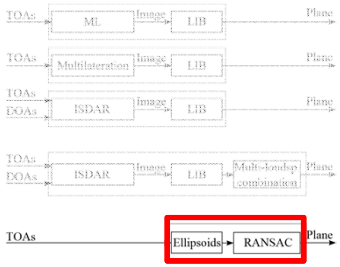


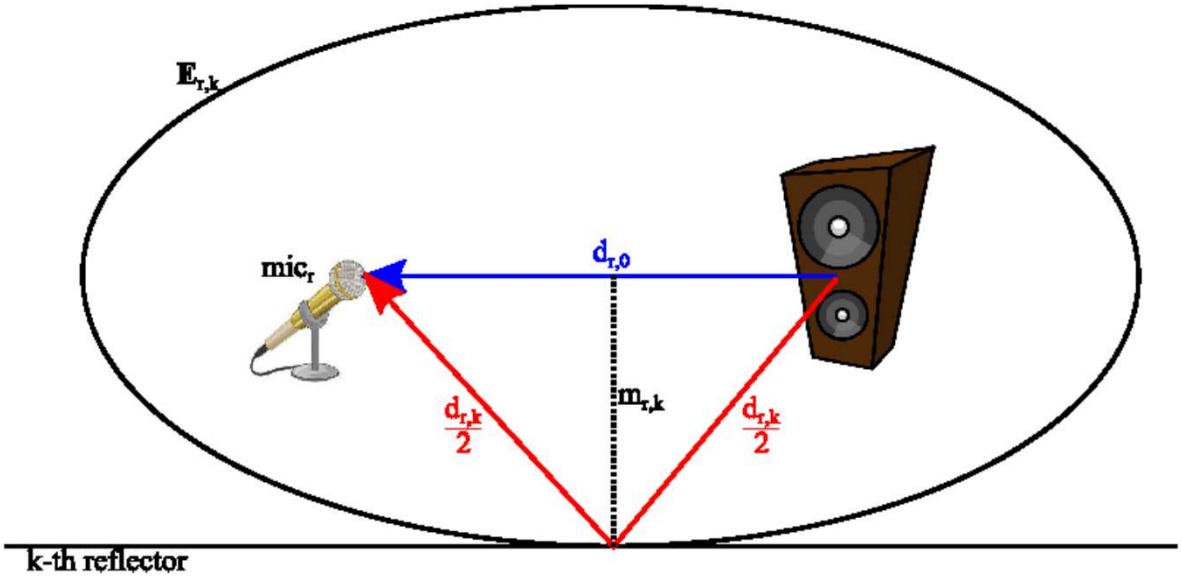
[1][Remaggi et al., IEEE TASLP, 2017]

Direct Localization

- Reflector localization
  - ETSAC



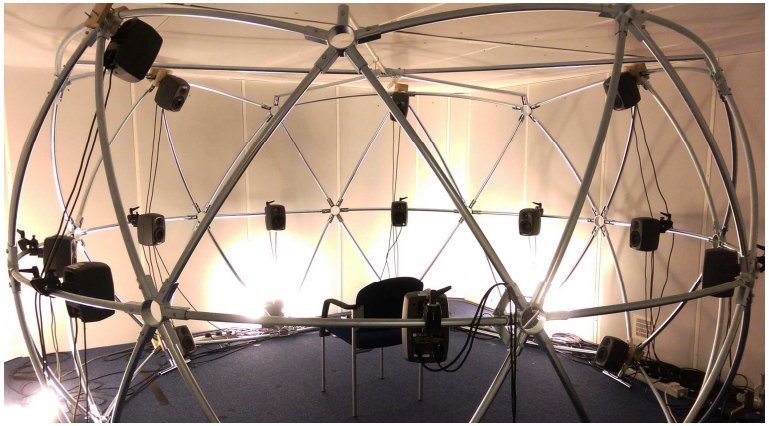




- Reflector search using RANSAC

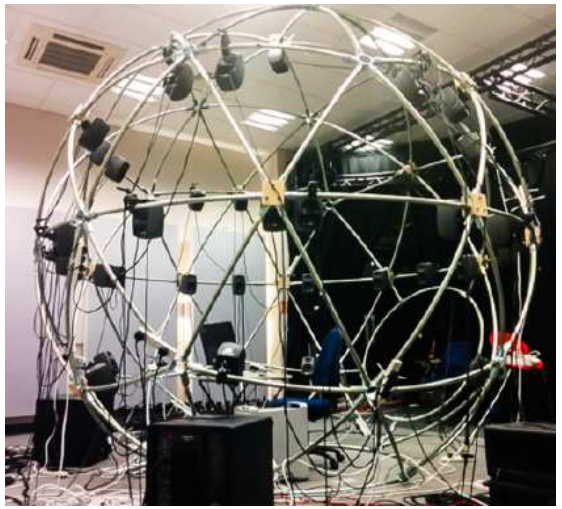
Experimental evaluation: the datasets

Studio1



AudioBooth

VML

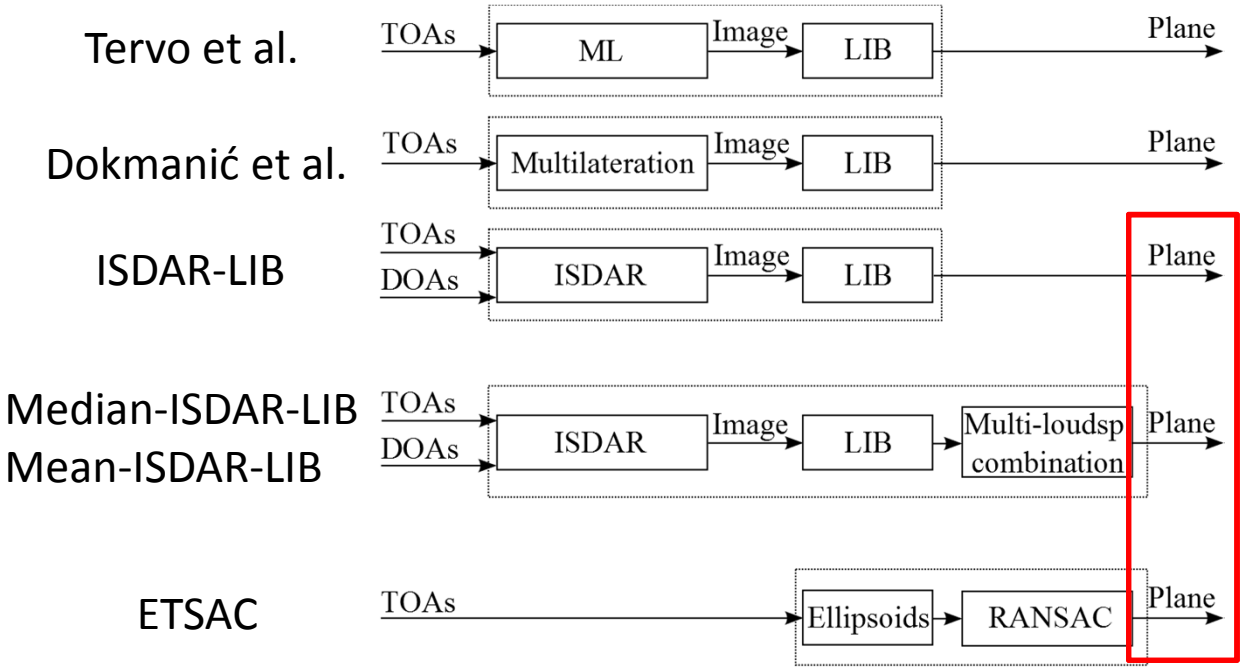


Vislab

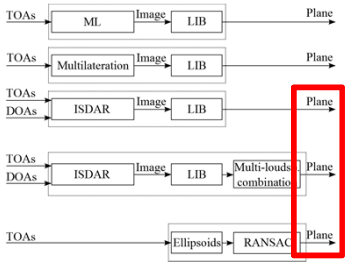
UNIVERSITY OF SURREY

Experimental evaluation: methods comparison

- Reflector localization
  - For each dataset average over all the combinations



Experimental evaluation: methods comparison

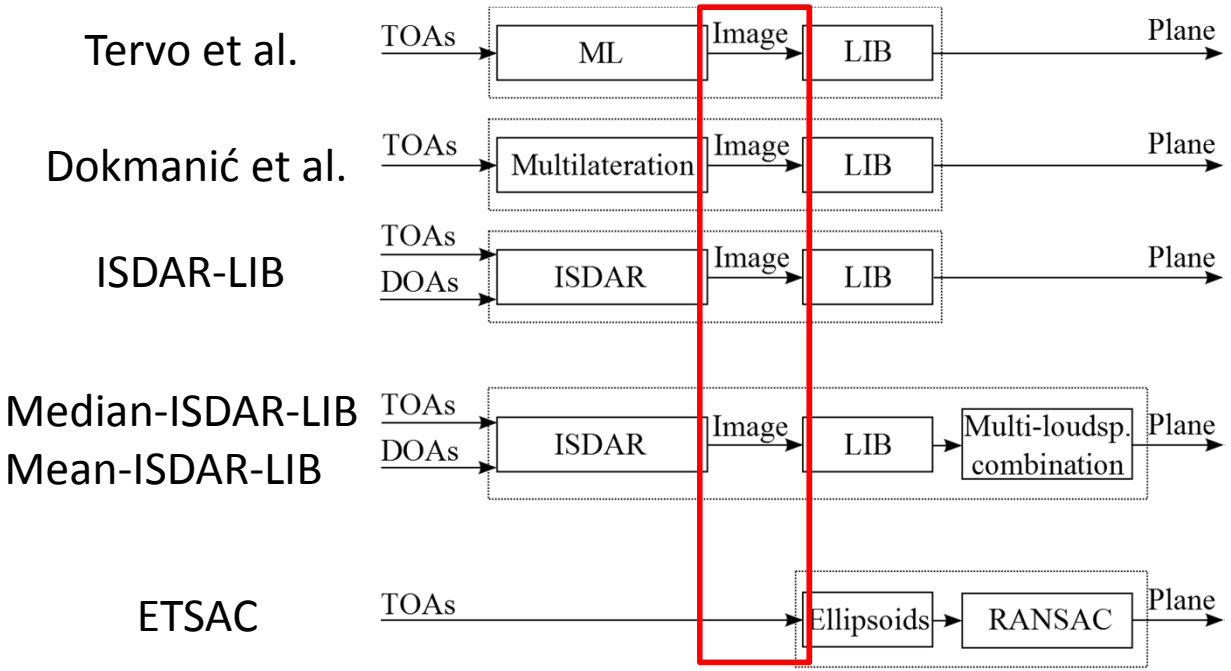


- Reflector localization

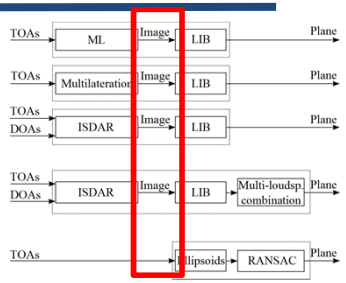
RMSE (mm)	AudioBooth	Vislab	VML	Studio1	AVG
ISDAR-LIB[1]	86	47	148	46	120 ± 20
Median-ISDAR-LIB[1]	92	70	120	54	96 ± 10
Mean_ISDAR-LIB[1]	56	59	127	49	90 ± 12
ETSAC[1]	**21**	**30**	**82**	**17**	**52 ± 2**

Experimental evaluation: methods comparison

- Image source localization
  - For each dataset average over all the combinations



Experimental evaluation: methods comparison



➢ Image source localization

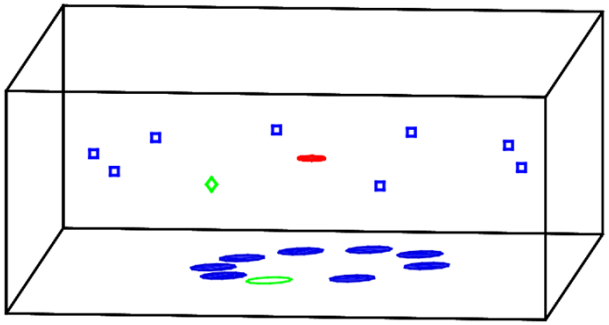
Gross errors (%)	AudioBooth	Vislab	VML	Studio1	AVG
Maximum-Likelihood[1]	67.7	70.3	89.0	66.0	73.3 ± 9.0
Multilateration[2]	18.8	25.8	100.0	5.8	37.6 ± 36.0
ISDAR-LIB[3]	**0.0**	**0.0**	68.2	**0.0**	17.1 ± 28.9
ETSAC[3]	**0.0**	**0.0**	**50.0**	**0.0**	**12.5 ± 21.2**
RMSE (mm)					
Maximum-Likelihood[1]	323	328	**342**	331	334 ± 6
Multilateration[2]	265	263	--	296	267 ± 10
ISDAR-LIB[3]	208	239	352	232	245 ± 4
ETSAC[3]	**82**	**163**	438	**100**	**220 ± 8**

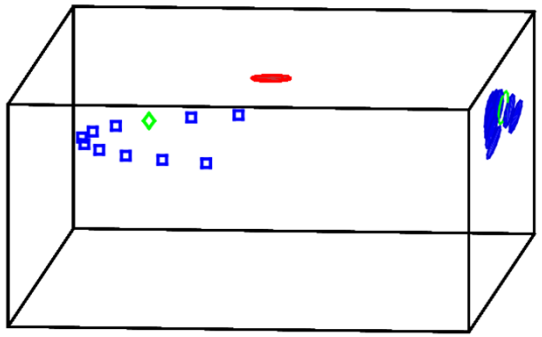
UNIVERSITY OF SURREY

[1]Tervo et al., ICASSP, 2012
[2]Dokmanić et al., PNAS, 2013
[3]Remaggi et al., IEEE TASLP, 2017

Conclusion

- Different acoustic reflector estimation methods have been presented
- ISDAR-LIB and two of its variants have been proposed for the *image source reversion* category
- ETSAC has been proposed for the *direct localization* category
- ETSAC is the best reflector localization method which has been tested





UNIVERSITY OF SURREY

Any question?

Please email Luca Remaggi at

l.remaggi@surrey.ac.uk

Thank you!

UNIVERSITY OF SURREY