Acoustic Reflector Localization and Classification

L. Remaggi[1], H. Kim[1], P. J. B. Jackson[1], F. Fazi[2] and A. Hilton[1]

[1]University of Surrey, UK
[2]University of Southampton, UK

ICASSP 2018 – 19th April
Calgary, Canada

Outline

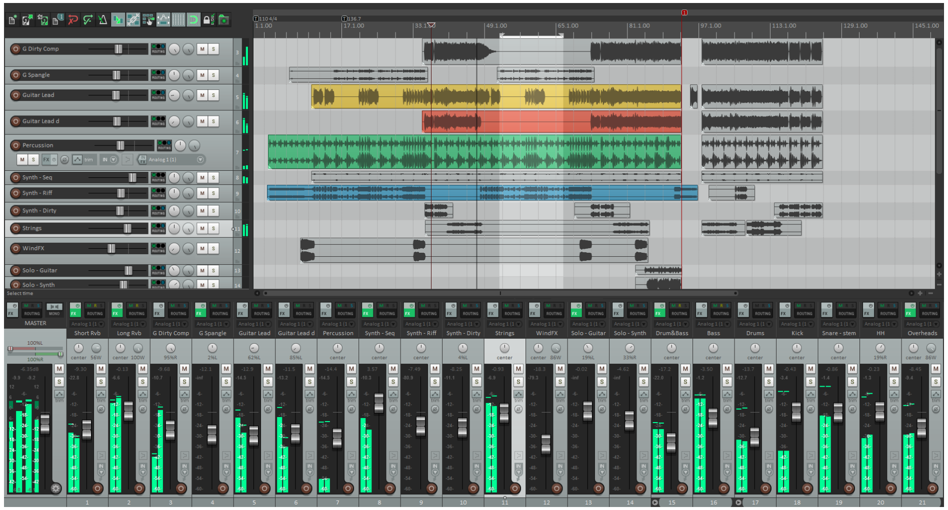
- Related Previous Work and Applications
- Acoustic Reflector Localization
- Acoustic Reflector Classification
- Experiments and Results
- Concusions

Applications

- Understanding of acoustic scenes targets several applications that span from audio forensics to source separation

Applications

- Understanding of acoustic scenes targets several applications that span from audio forensics to source separation
- Here, the two main targeted areas are active manipulation of **reverberant spatial audio objects** and **mixed reality**

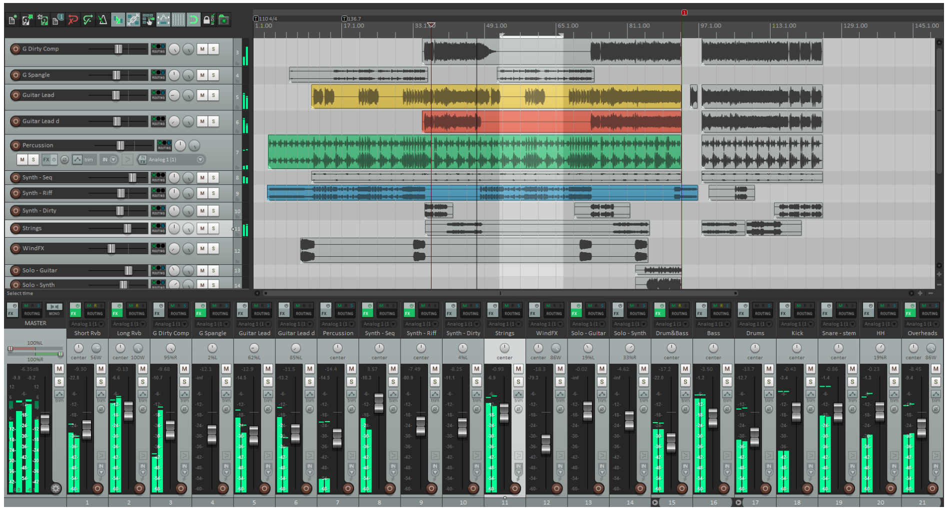


Applications

- Understanding of acoustic scenes targets several applications that span from audio forensics to source separation

- Here, the two main targeted areas are active manipulation of **reverberant spatial audio objects** and **mixed reality**

- It is important to identify not only the position of the acoustic reflectors but also their size, in particular for **non-static listener scenarios**



Related Work – Acoustic Scene Analysis from RIRs

Acoustic Reflector Localization

Acoustic Reflector Classification

Related Work – Acoustic Scene Analysis from RIRs

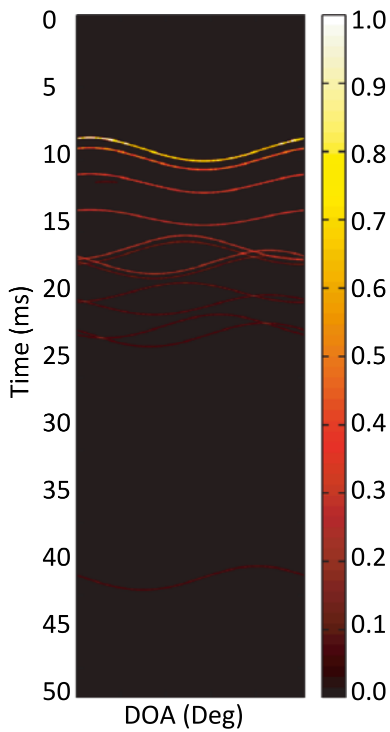
Acoustic Reflector Localization

- Methods using the **space-time domain**[Kuster et al., *JASA*, 2004]
- Geometric approaches using the reflection **time difference of arrival (TDOA)** [Antonacci et al., IEEE TASLP, 2012]
- Geometric approaches joining the reflection **time of arrival (TOA) with direction of arrival (DOA)** [Remaggi et al., IEEE TASLP, 2017].

Acoustic Reflector Classification

Related Work – Acoustic Scene Analysis from RIRs

Acoustic Reflector Localization

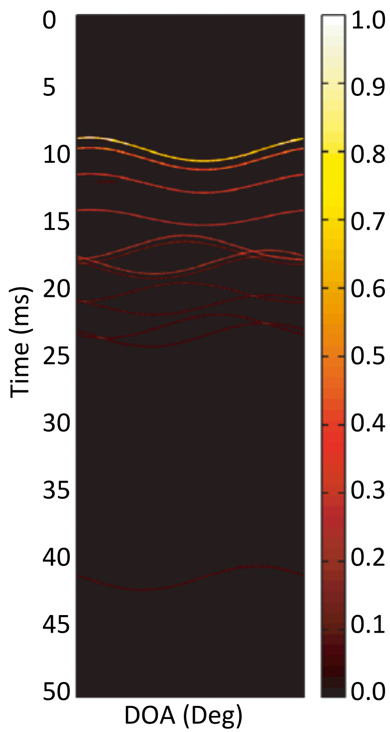


- Methods employing the **DOA-time domain** of room impulse responses [Torres et al., JASA, 2013]

Acoustic Reflector Classification

Related Work – Acoustic Scene Analysis from RIRs

Acoustic Reflector Localization



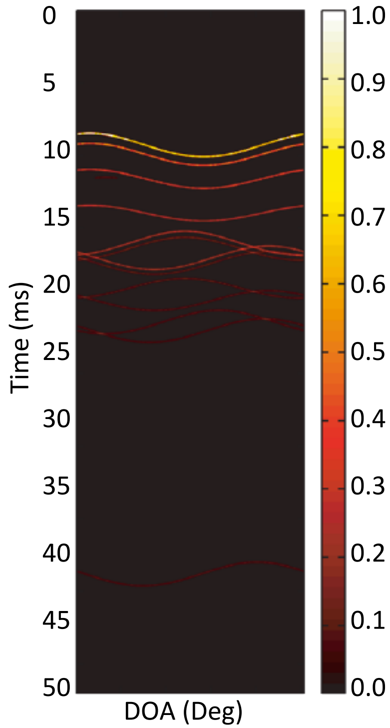
- Methods employing the **DOA-time domain** of room impulse responses [Torres et al., JASA, 2013]

Acoustic Reflector Classification

- There is still no method to classify reflectors in terms of their size

Related Work – Acoustic Scene Analysis from RIRs

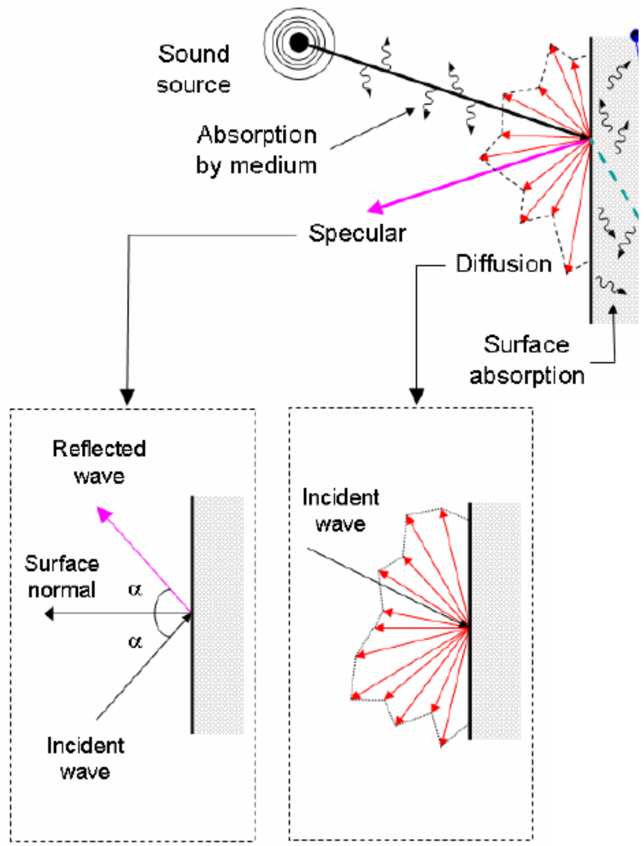
Acoustic Reflector Localization



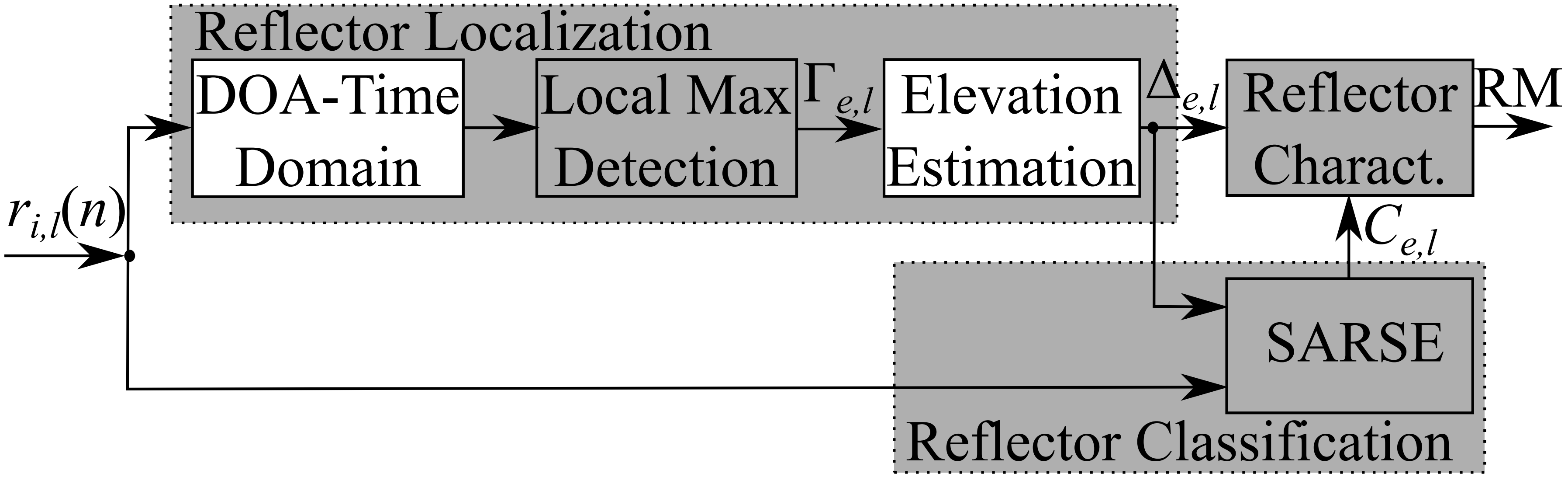
- Methods employing the **DOA-time domain** of room impulse responses [Torres et al., JASA, 2013]

Acoustic Reflector Classification

- There is still no method to classify reflectors in terms of their size

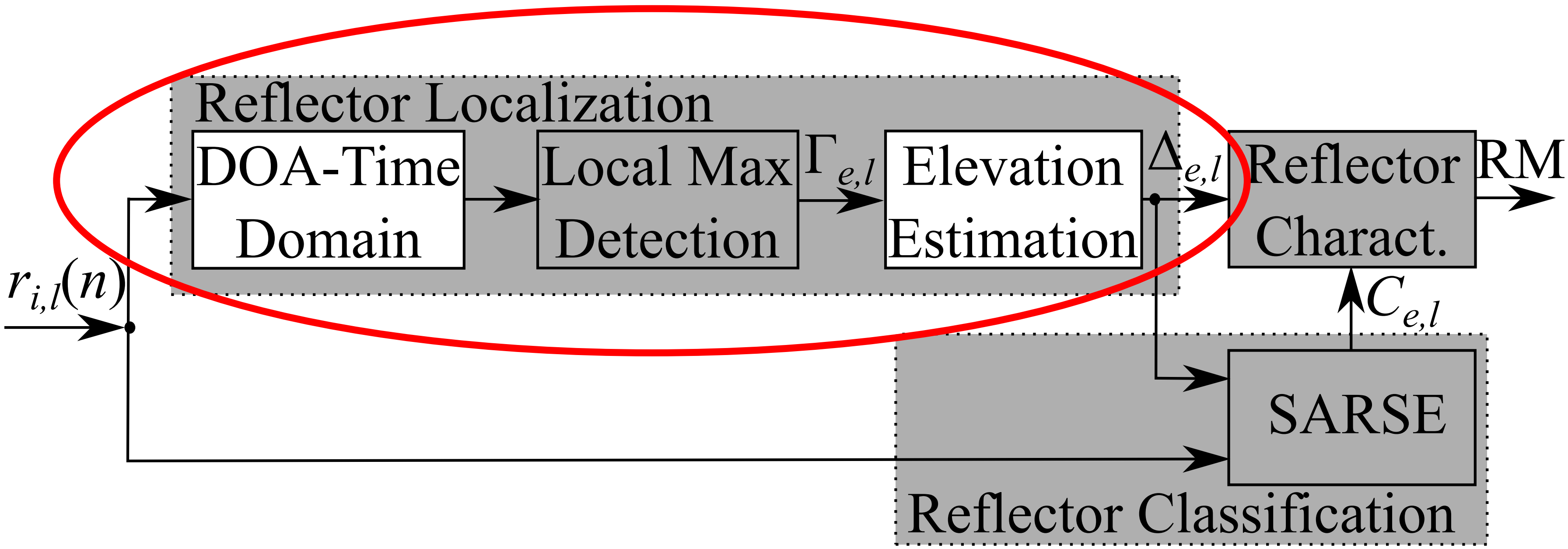


Acoustic Reflector Localization and Classification



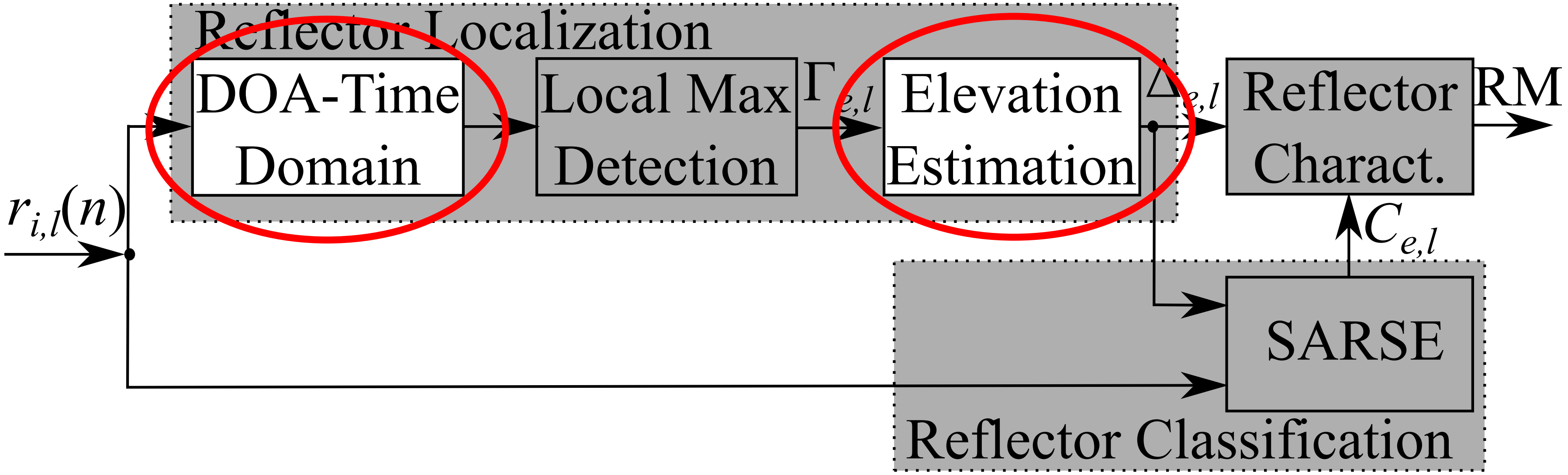
Acoustic Reflector Localization and Classification

- Acoustic reflector localization based on DOA-time domain analysis



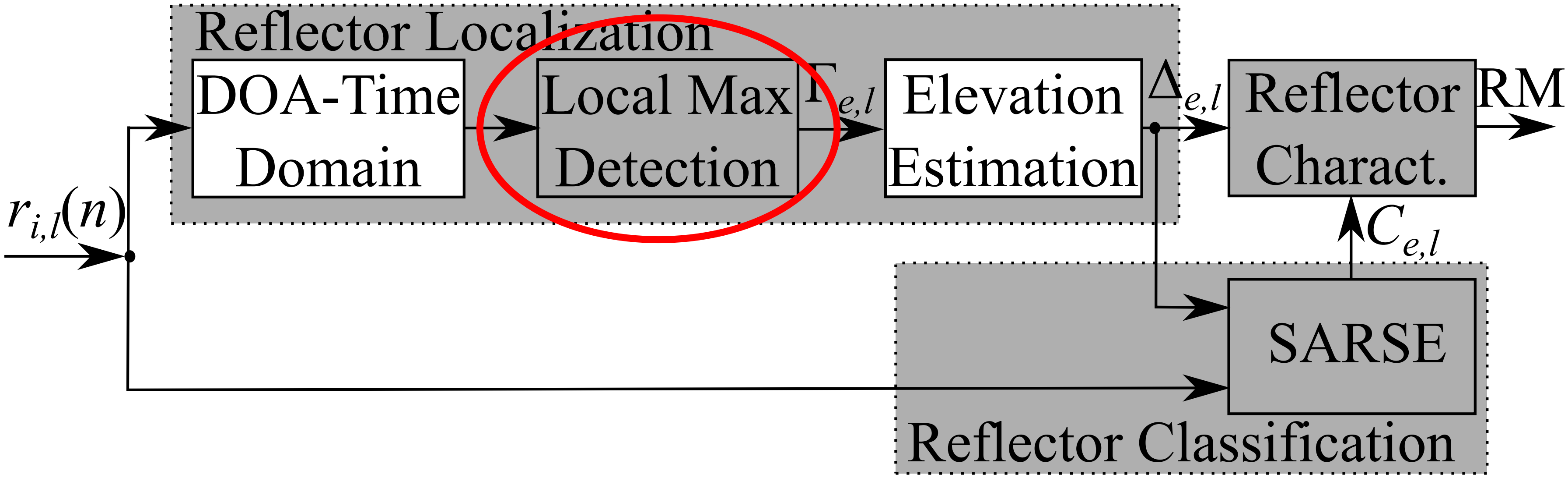
Acoustic Reflector Localization and Classification

- Acoustic reflector localization based on DOA-time domain analysis
- Audio processing

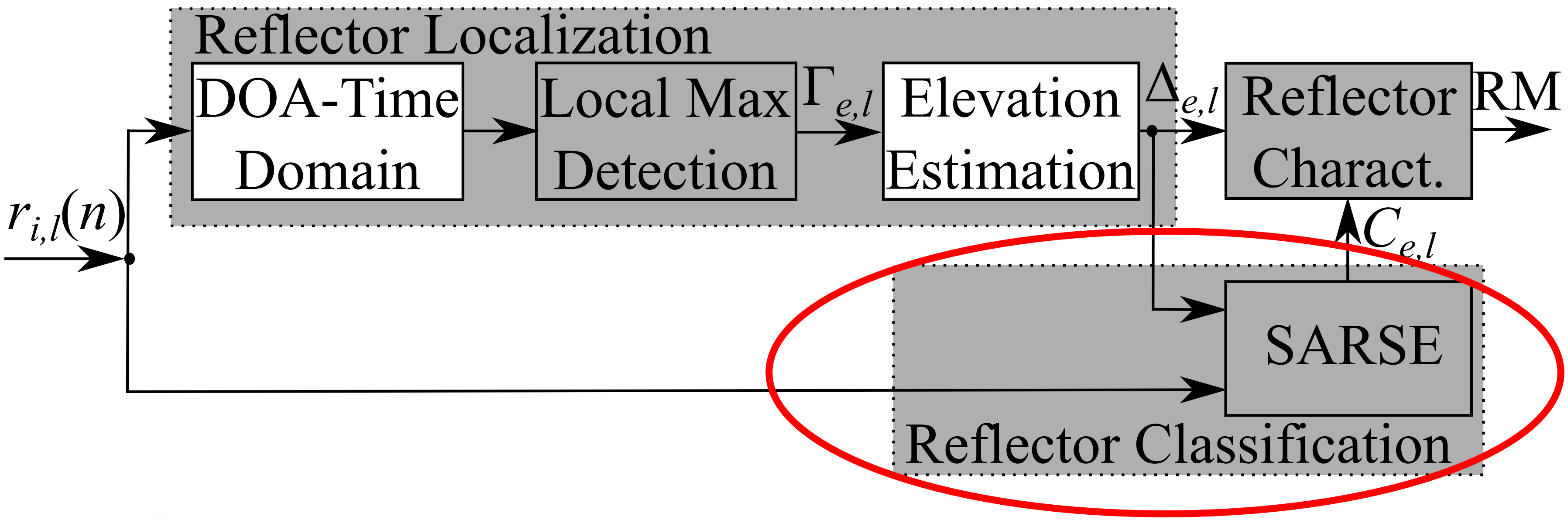


Acoustic Reflector Localization and Classification

- Acoustic reflector localization based on DOA-time domain analysis
- Audio processing
- Image processing

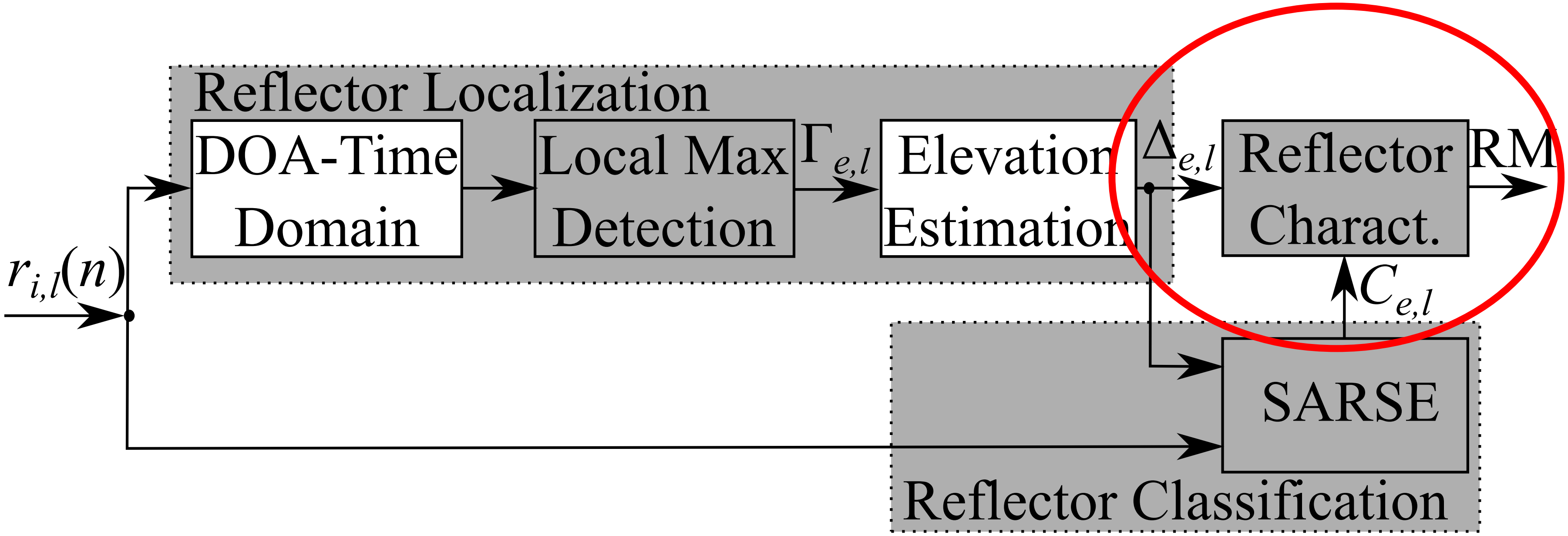


Acoustic Reflector Localization and Classification

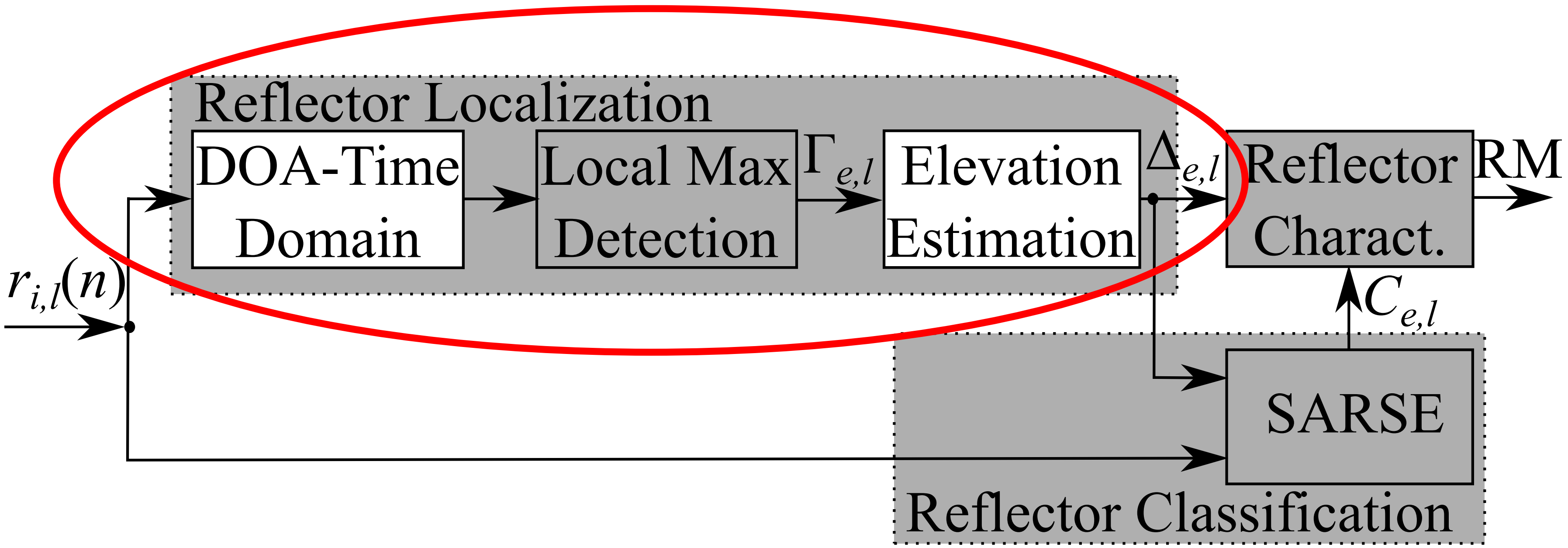


- Acoustic reflector classification algorithm named as Scattering Analysis based Reflector Size Estimator (SARSE)

Acoustic Reflector Localization and Classification



Acoustic Reflector Localization and Classification

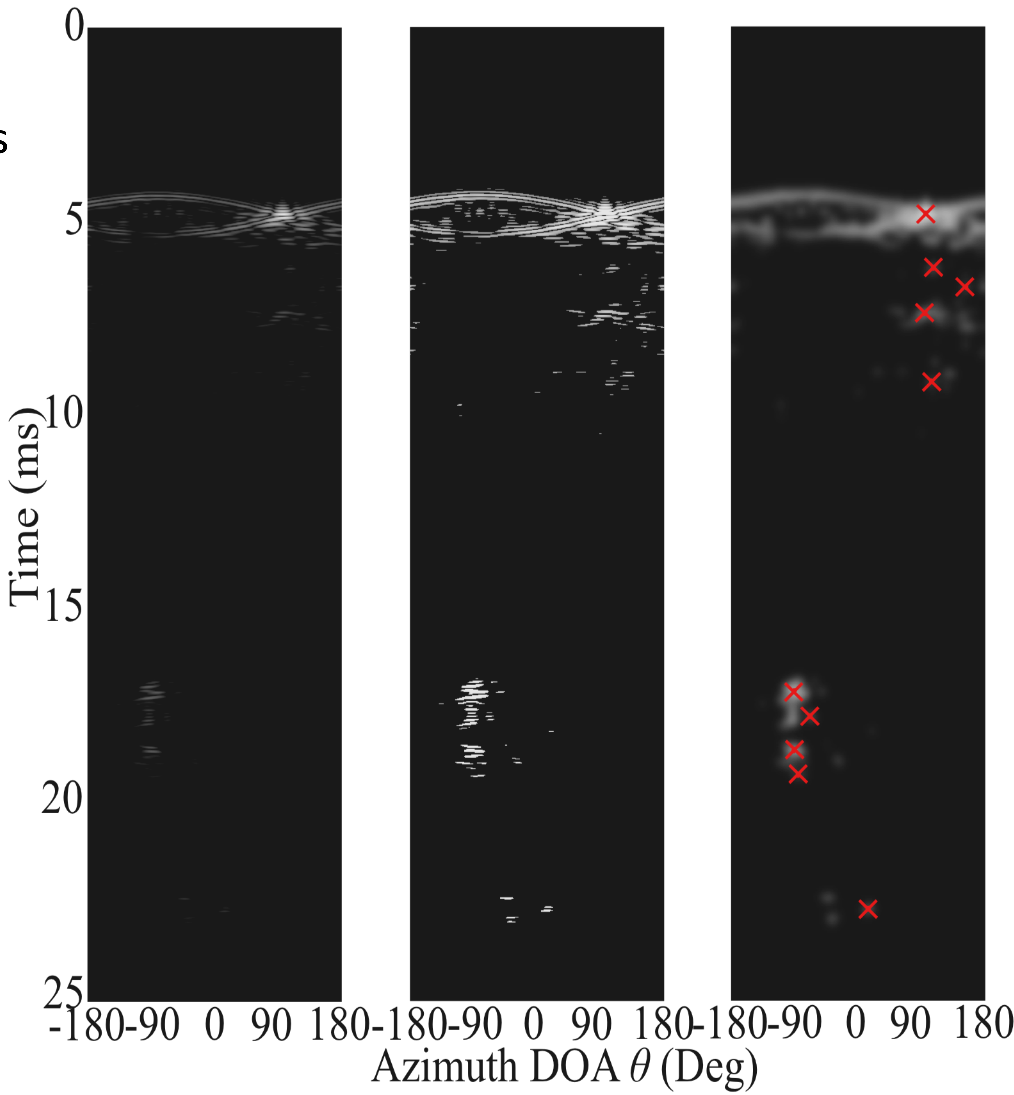


Acoustic Reflector Localization

- A **superdirective array beamformer** was used to steer the beam towards every direction in 3D [Kim et al., 3DV Conference, 2017]

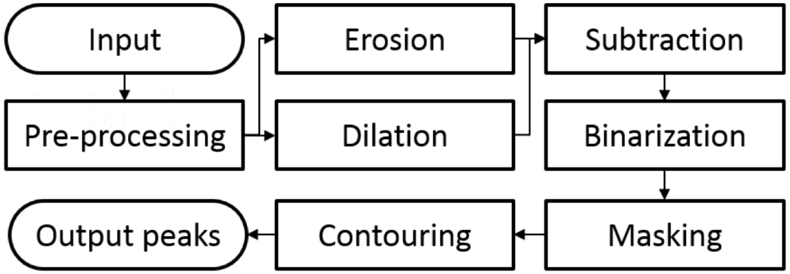
Acoustic Reflector Localization

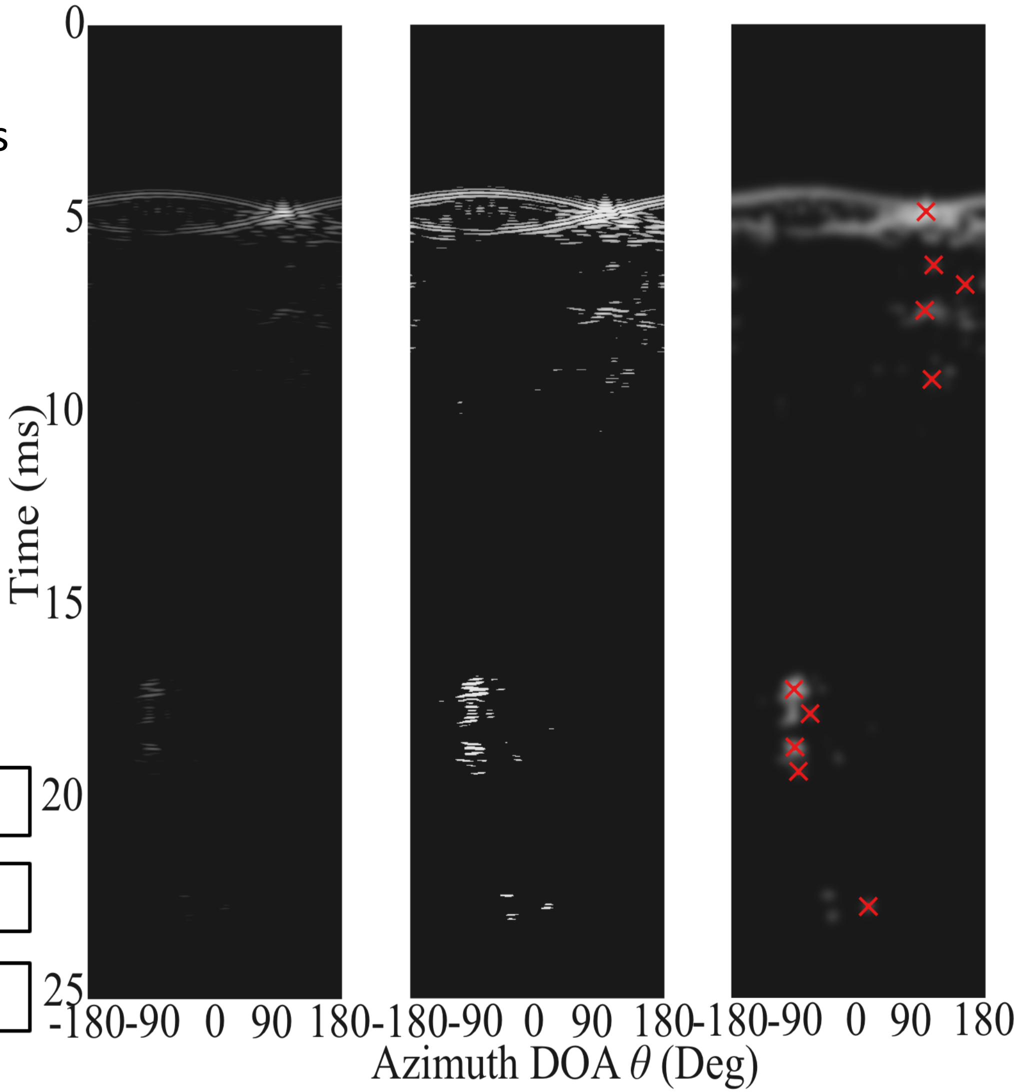
- A **superdirective array beamformer** was used to steer the beam towards every direction in 3D [Kim et al., 3DV Conference, 2017]
- By fixing the elevation to 0° the beamformed RIRs can be visualized
- **One loudspeaker** and **circular array**



Acoustic Reflector Localization

- A **superdirective array beamformer** was used to steer the beam towards every direction in 3D [Kim et al., 3DV Conference, 2017]
- By fixing the elevation to 0° the beamformed RIRs can be visualized
- **One loudspeaker** and **circular array**
- **Image processing** methods were applied to find the maxima, thus, estimate the reflection TOAs and DOAs

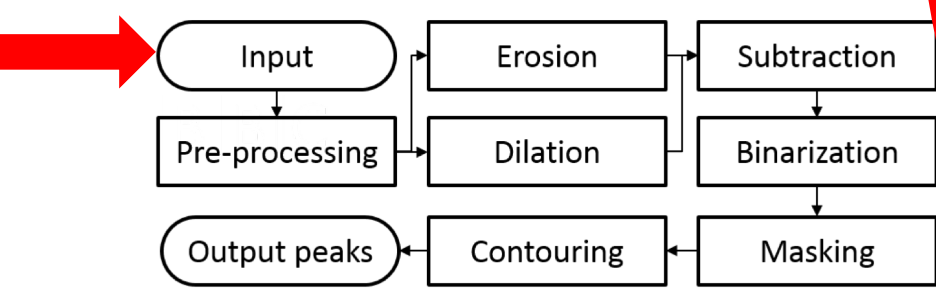


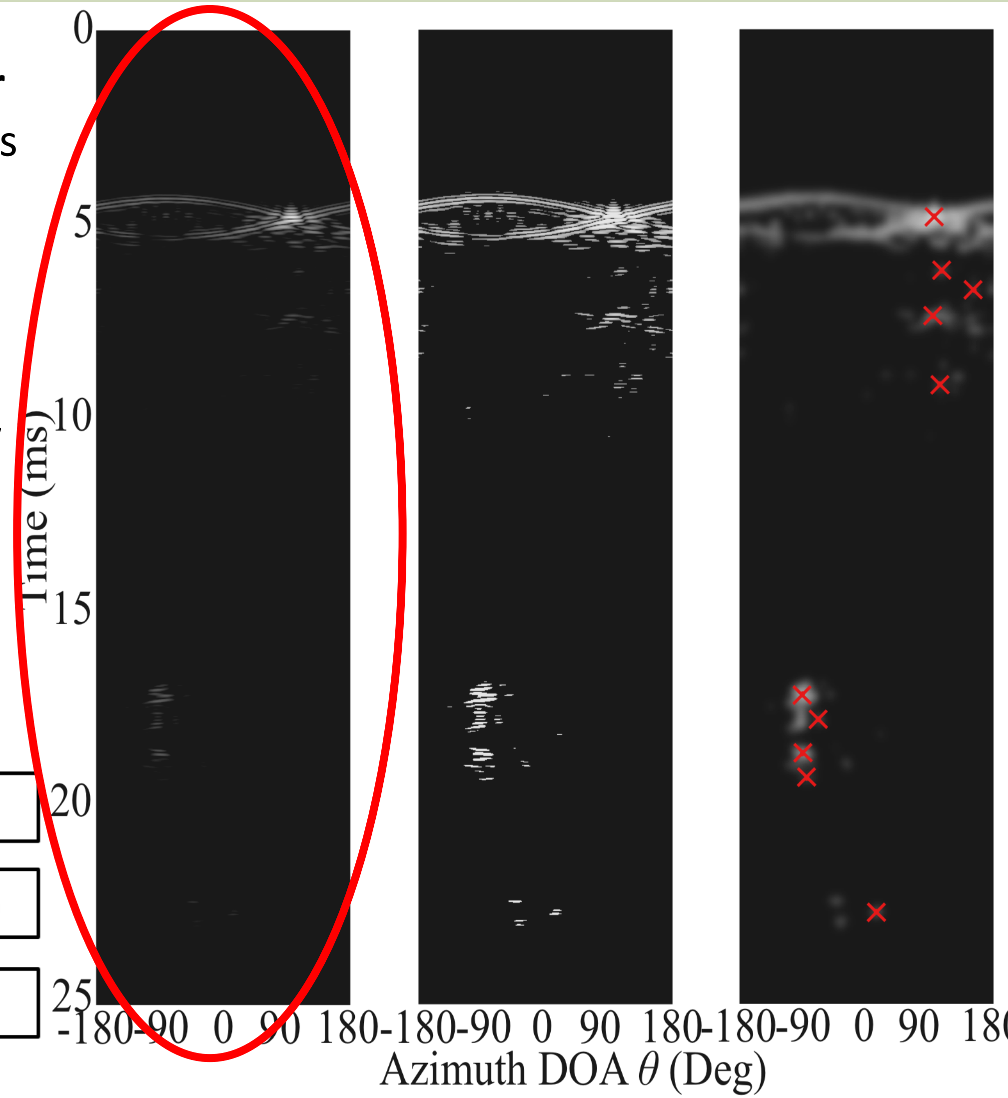


Acoustic Reflector Localization

- A **superdirective array beamformer** was used to steer the beam towards every direction in 3D [Kim et al., 3DV Conference, 2017]
- By fixing the elevation to 0º the beamformed RIRs can be visualized
- **One loudspeaker** and **circular array**

- **Image processing** methods were applied to find the maxima, thus, estimate the reflection TOAs and DOAs

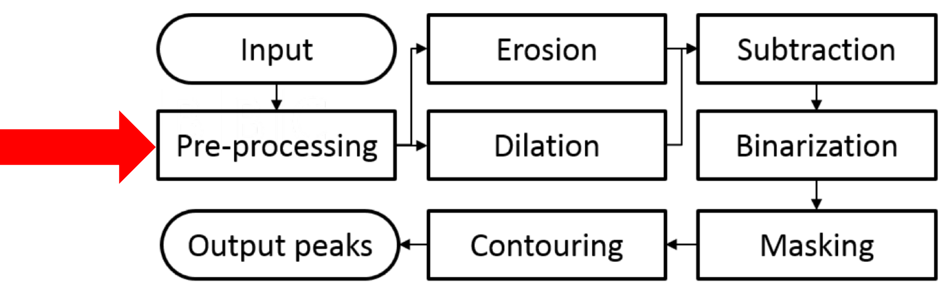


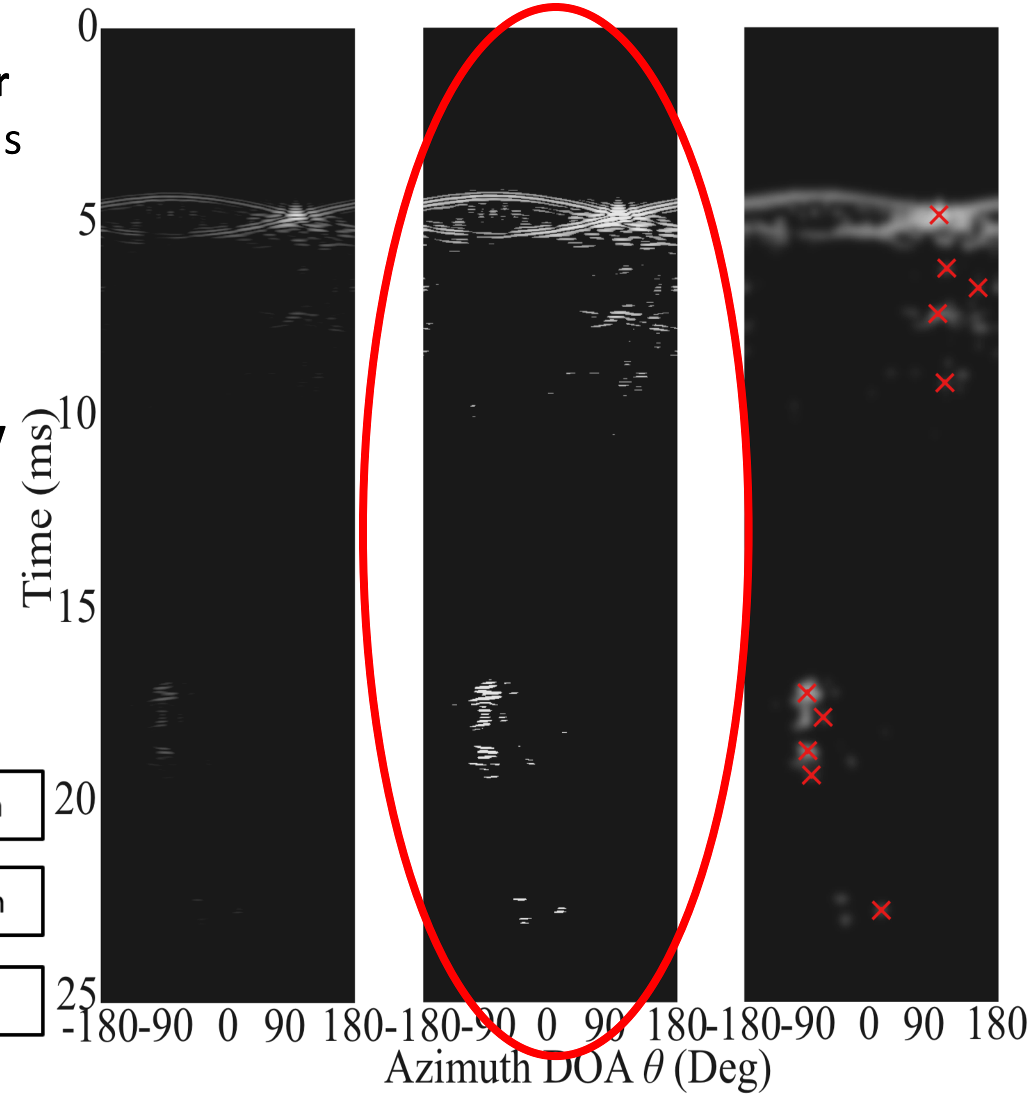


Acoustic Reflector Localization

- A **superdirective array beamformer** was used to steer the beam towards every direction in 3D [Kim et al., 3DV Conference, 2017]
- By fixing the elevation to 0º the beamformed RIRs can be visualized
- **One loudspeaker** and **circular array**

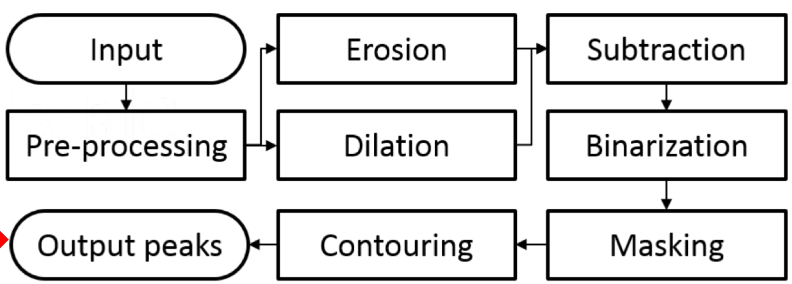
- **Image processing** methods were applied to find the maxima, thus, estimate the reflection TOAs and DOAs

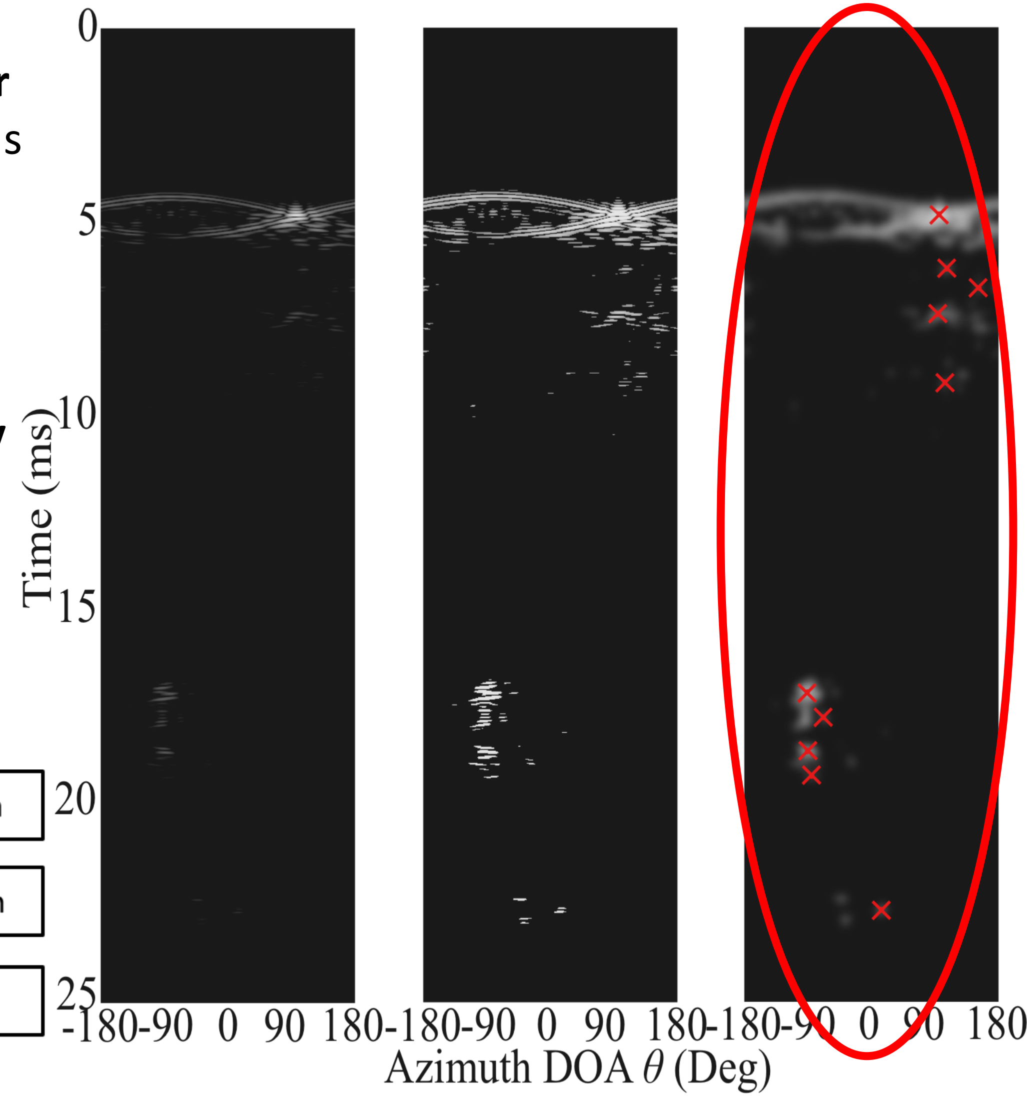




Acoustic Reflector Localization

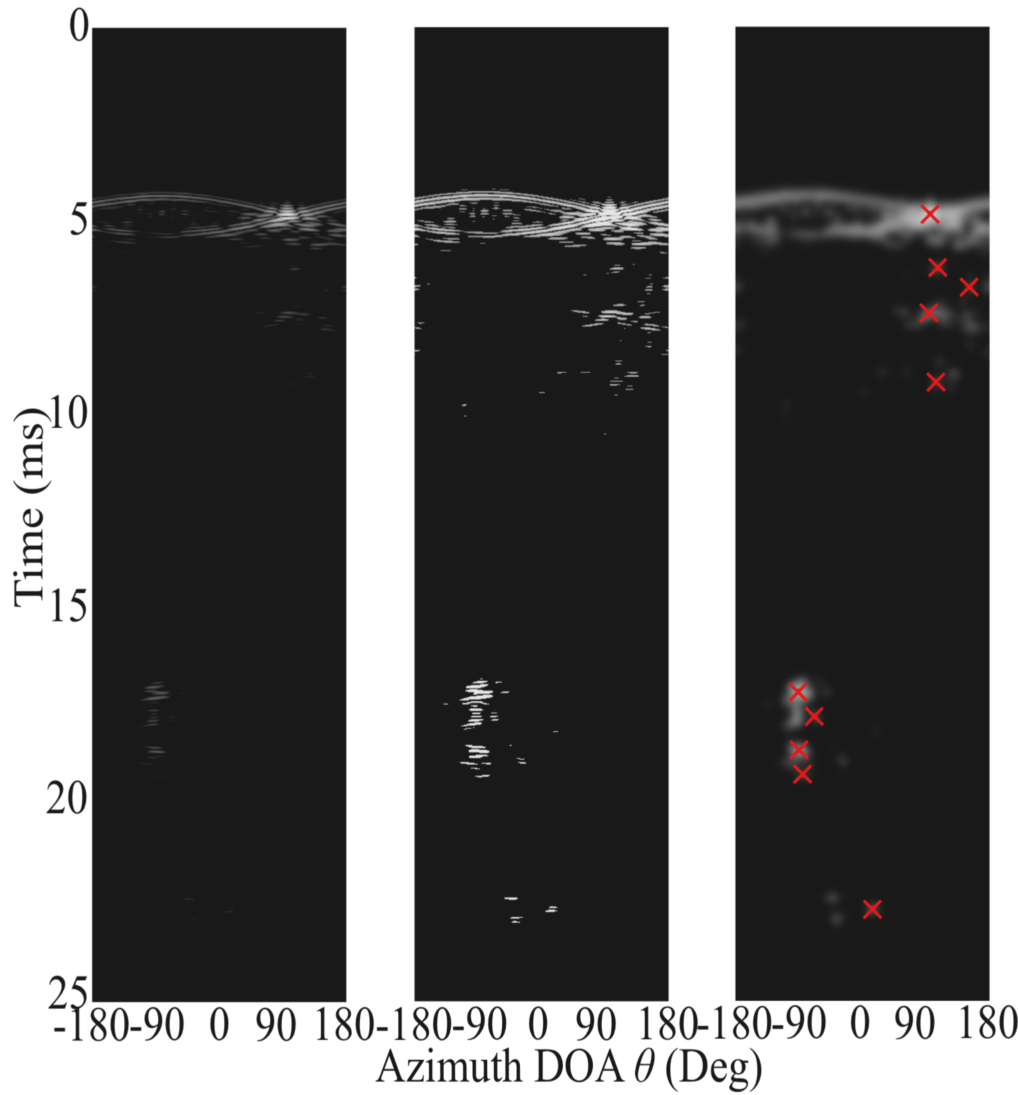
- A **superdirective array beamformer** was used to steer the beam towards every direction in 3D [Kim et al., 3DV Conference, 2017]
- By fixing the elevation to 0º the beamformed RIRs can be visualized
- **One loudspeaker** and **circular array**
- **Image processing** methods were applied to find the maxima, thus, estimate the reflection TOAs and DOAs





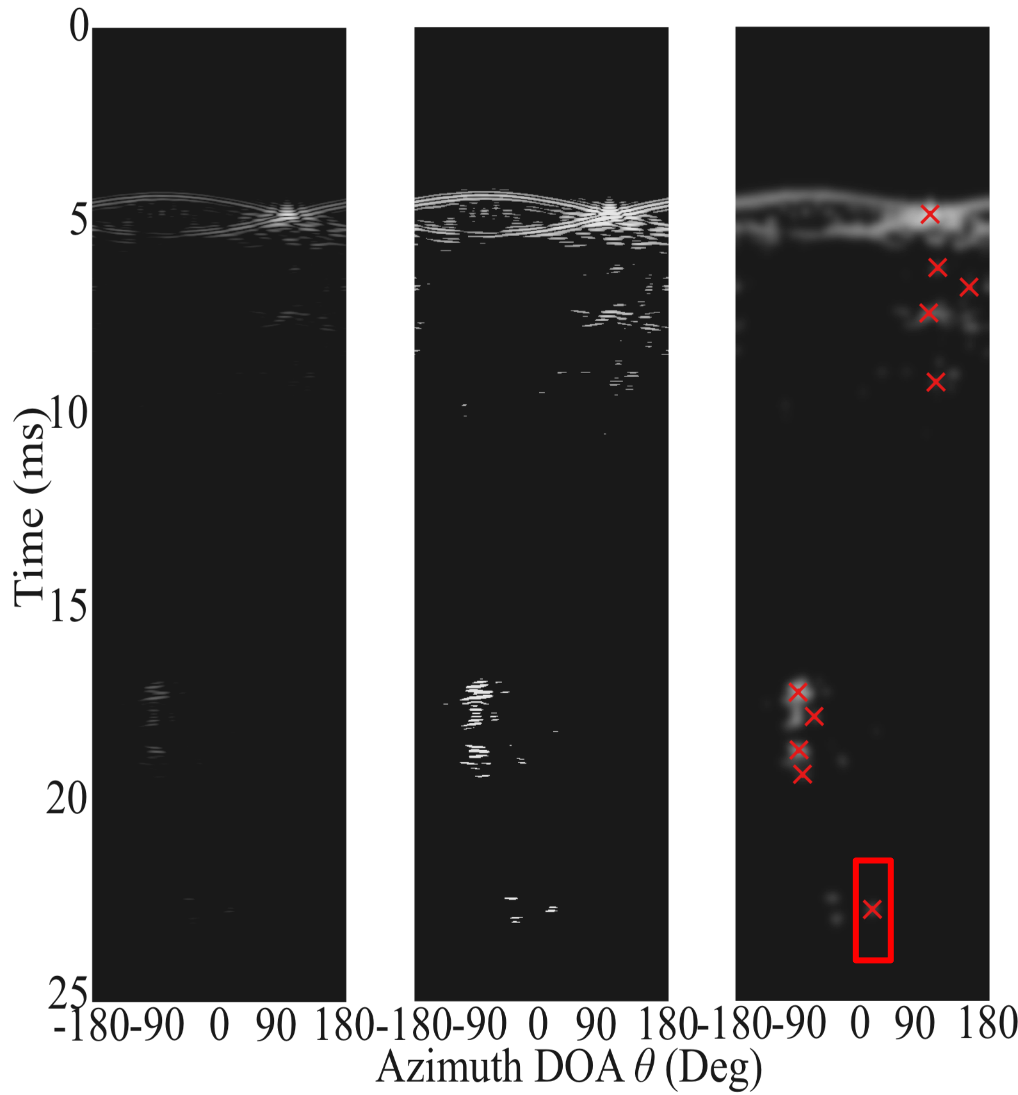
Acoustic Reflector Localization

- By knowing the TOA and azimuth DOA of the reflections, the segments are analyzed looking at their elevations, and the elevation DOA estimated as the angle providing the maximum signal energy

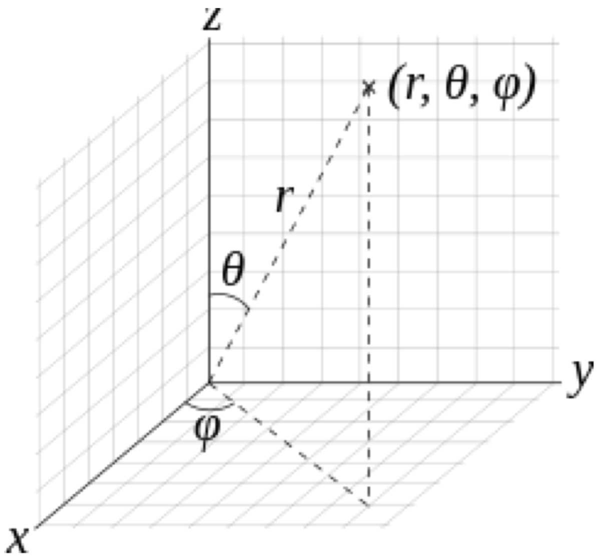


Acoustic Reflector Localization

- By knowing the TOA and azimuth DOA of the reflections, the segments are analyzed looking at their elevations, and the elevation DOA estimated as the angle providing the maximum signal energy

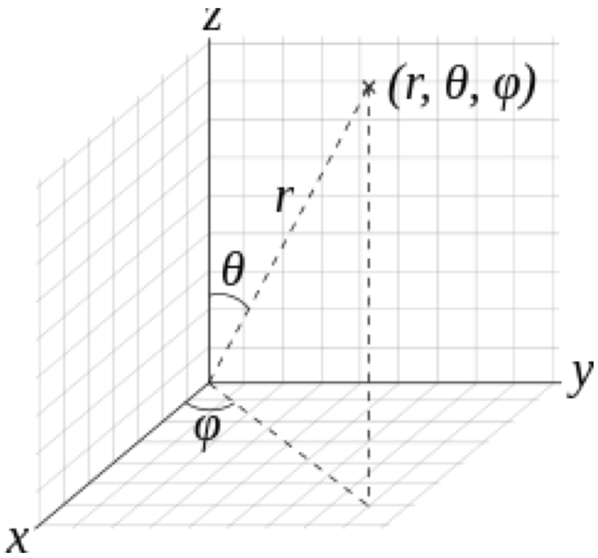


Acoustic Reflector Localization



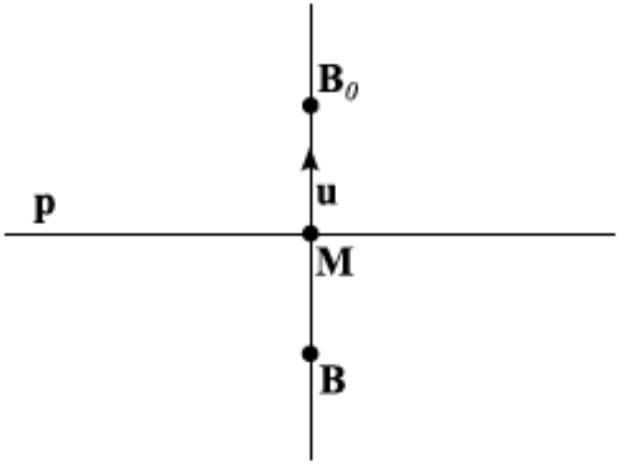
- Similar to Image source direction and ranging (ISDAR) [Remaggi et al., IEEE TASLP, 2017]

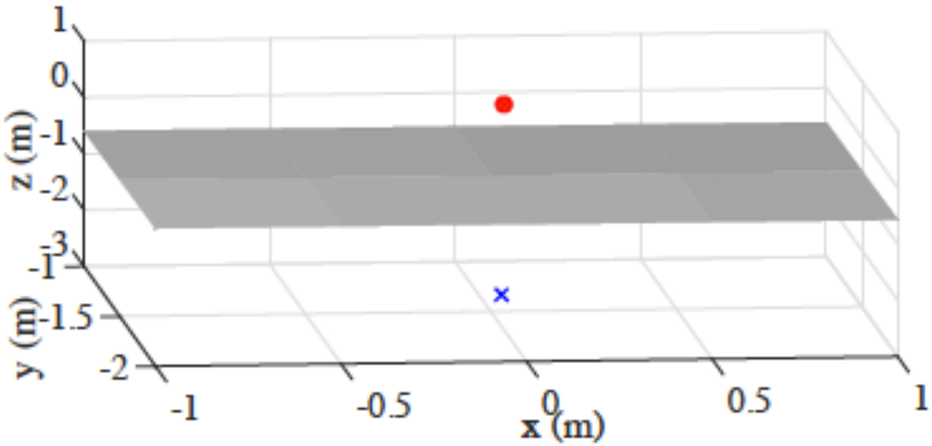
Acoustic Reflector Localization



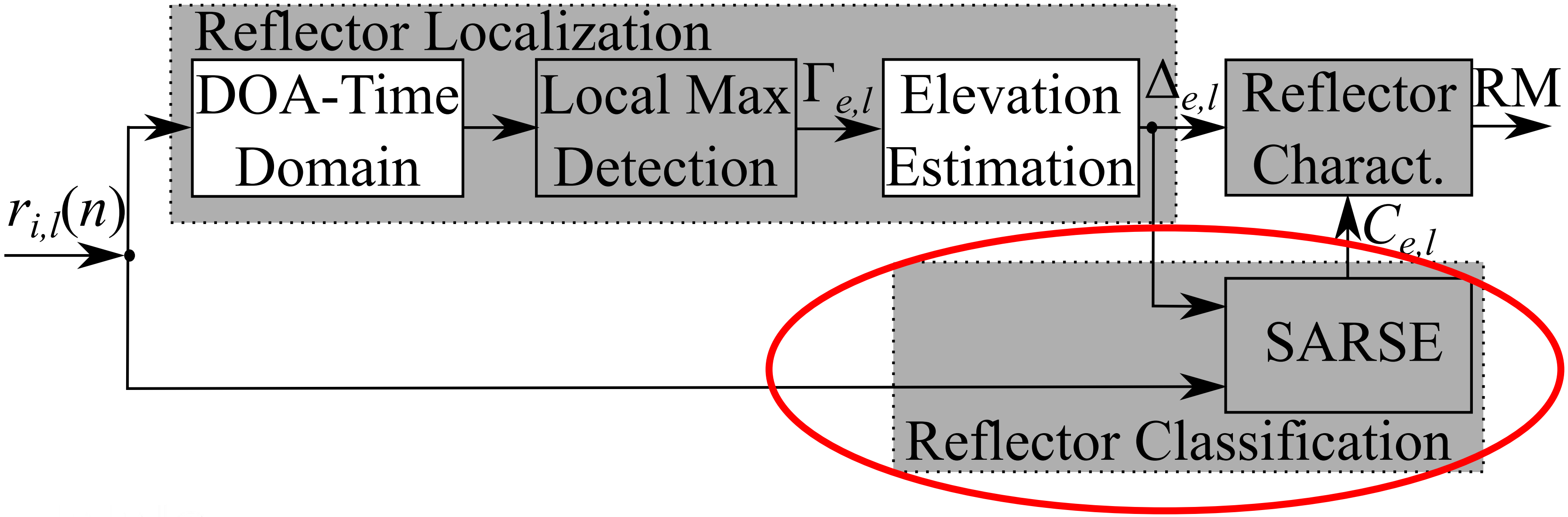
- Similar to Image source direction and ranging (ISDAR) [Remaggi et al., IEEE TASLP, 2017]

- Loudspeaker-image source bisection algorithm [Dokmanić et al., *PNAS*, 2014]

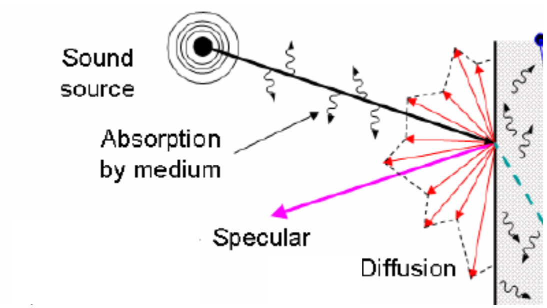




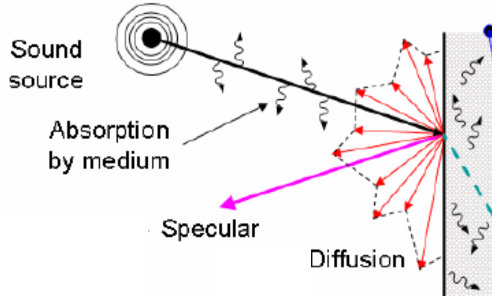
Acoustic Reflector Localization and Classification



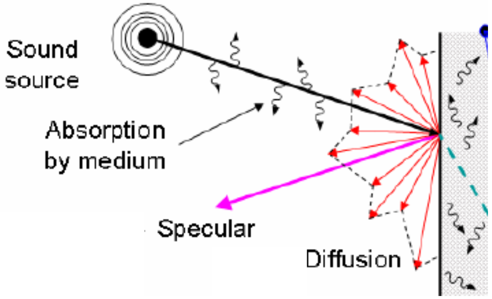
Acoustic Reflector Classification – The SARSE Algorithm



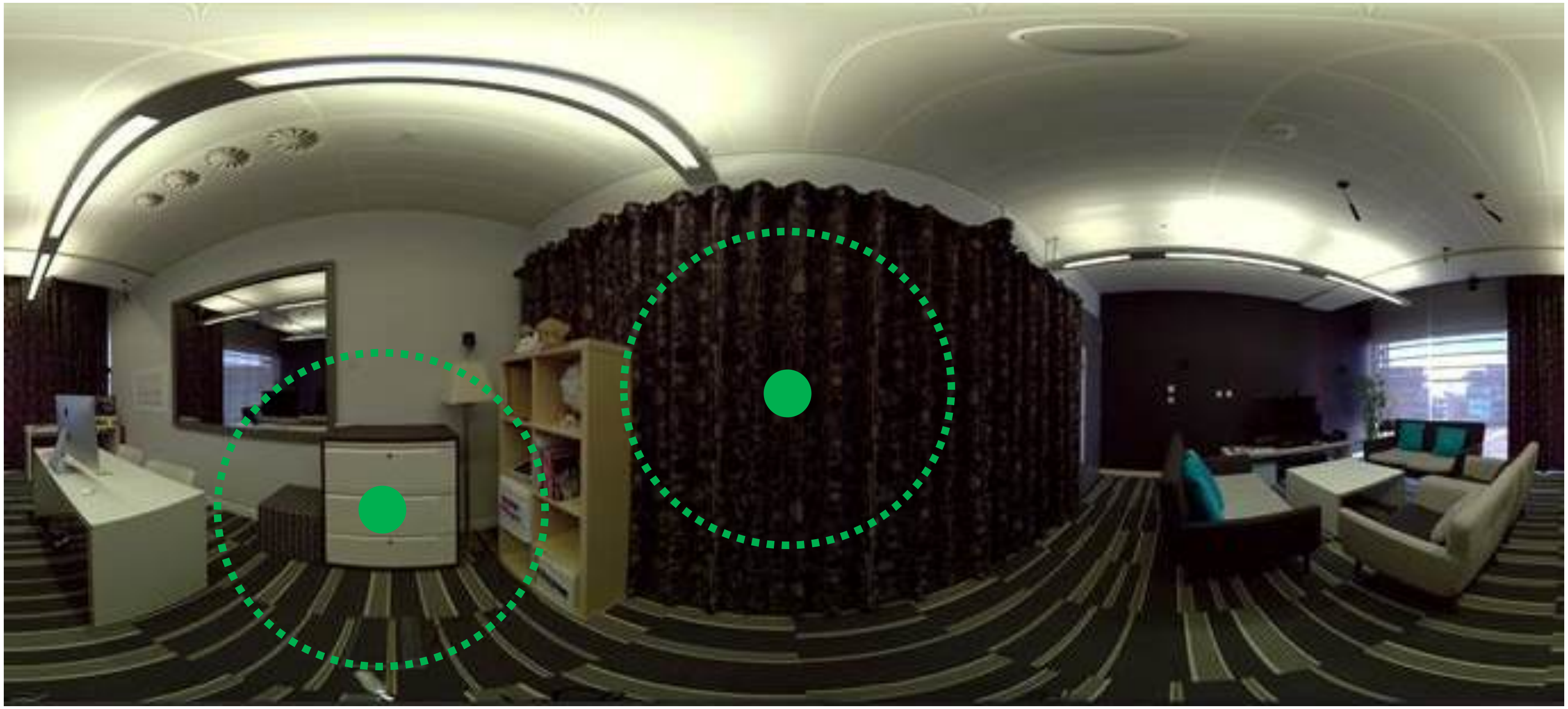
Acoustic Reflector Classification – The SARSE Algorithm

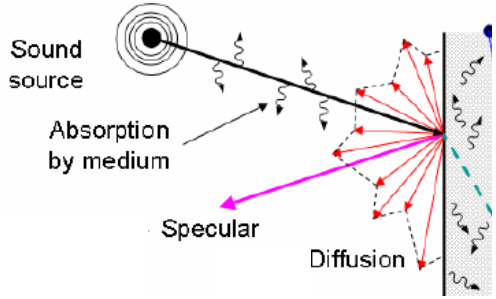


Acoustic Reflector Classification – The SARSE Algorithm

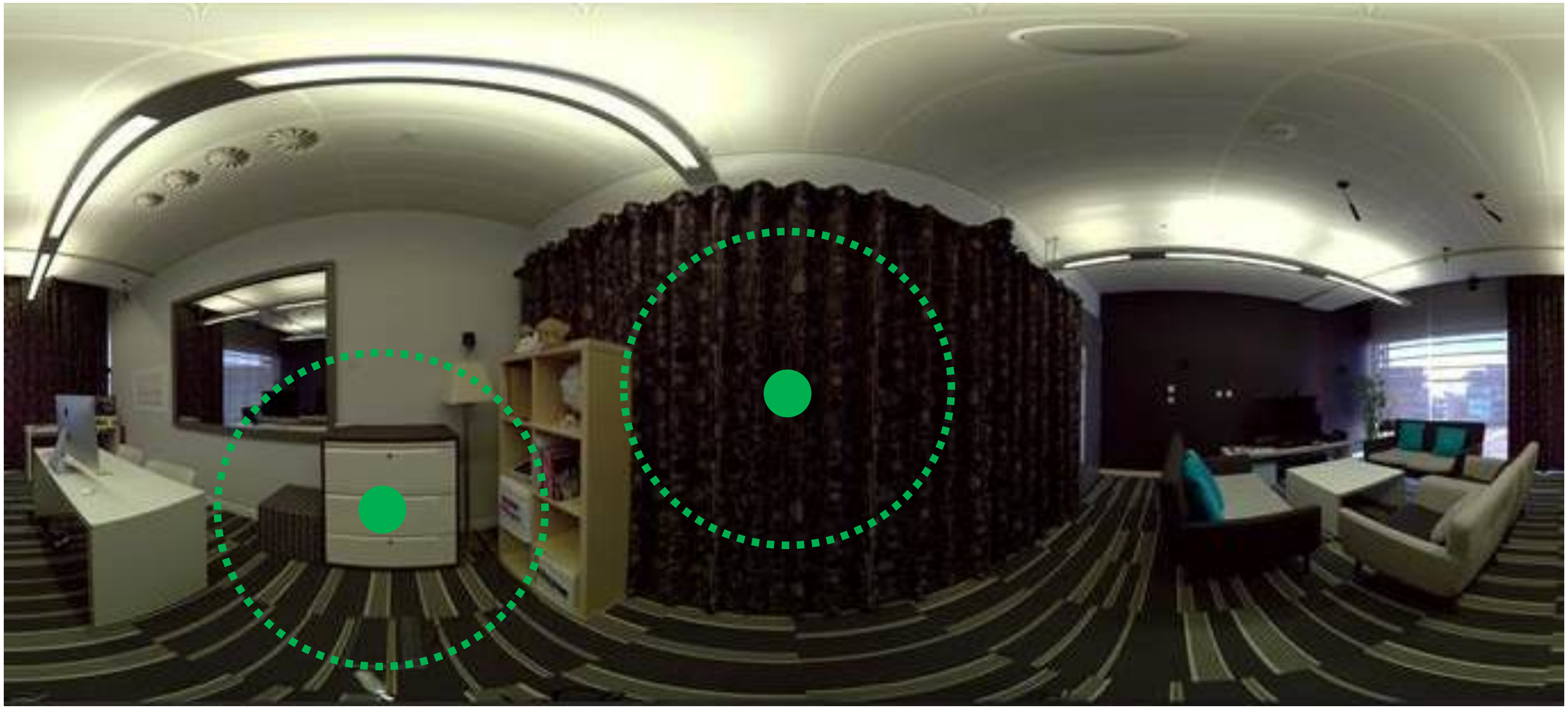


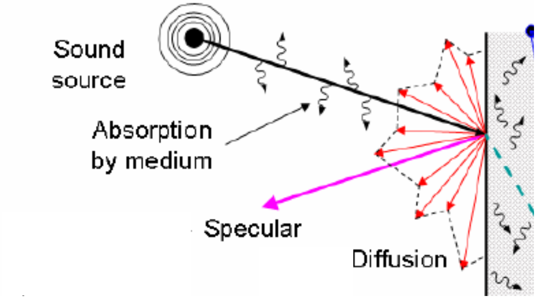
Acoustic Reflector Classification – The SARSE Algorithm





Acoustic Reflector Classification – The SARSE Algorithm

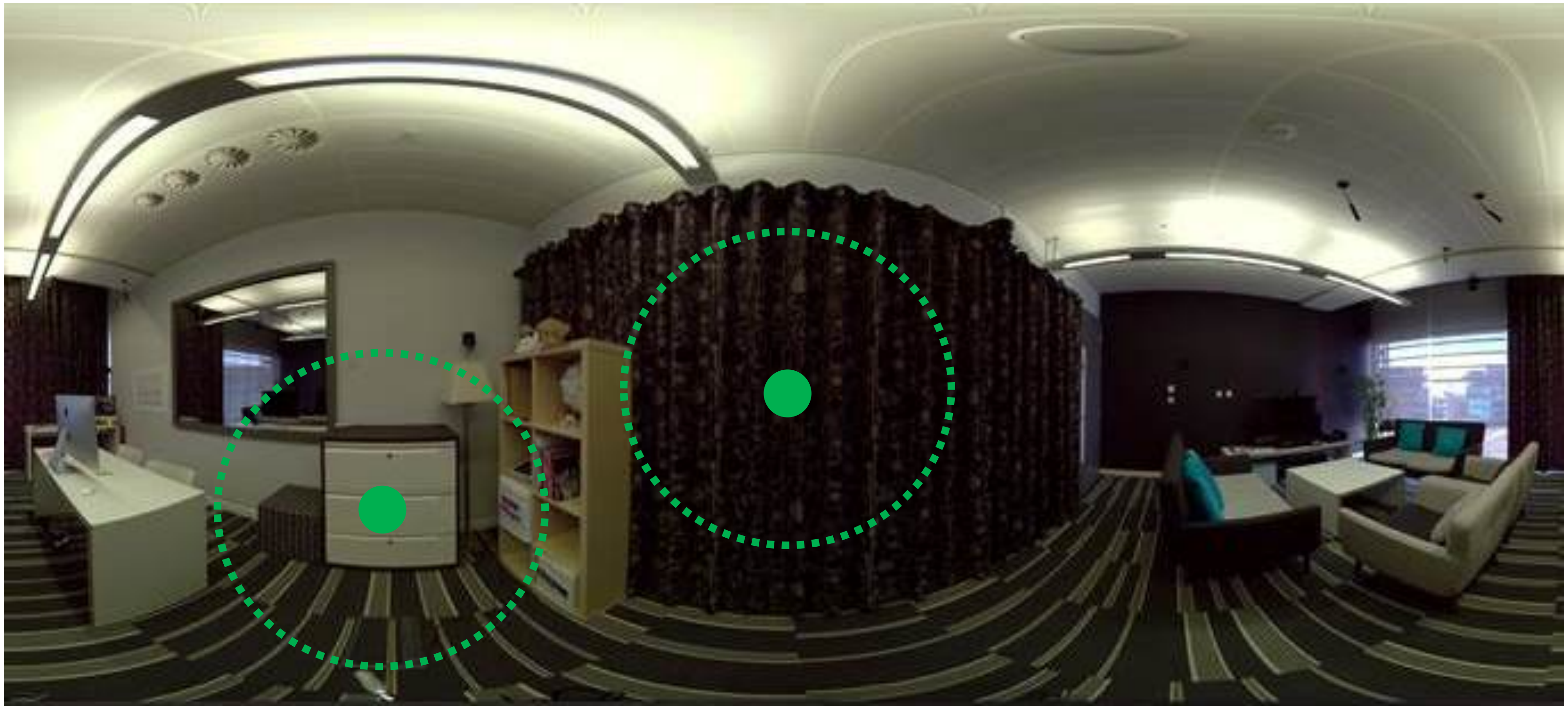


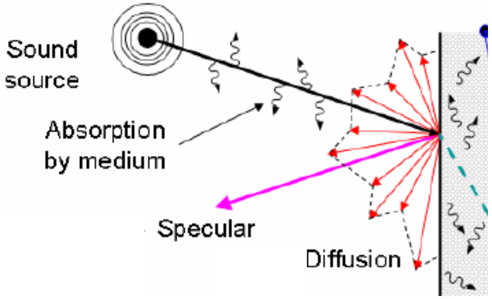


$$K = 1 - \frac{E^{sp}}{E^{sc}}$$

Partial reflection
Scattering coefficient

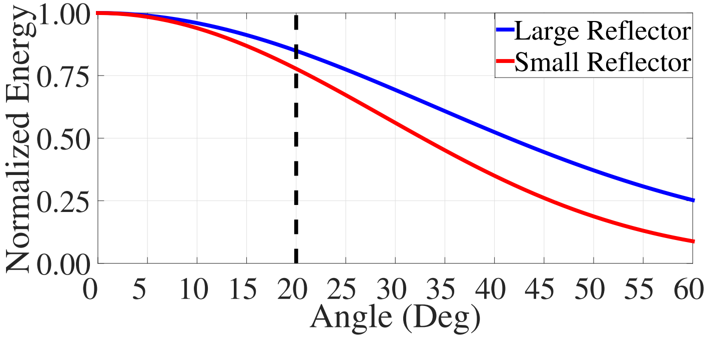
Acoustic Reflector Classification – The SARSE Algorithm





$$K = 1 - \frac{E^{sp}}{E^{sc}}$$

Partial reflection
Scattering coefficient



Experiments and Results

Classification

- Considering planar reflectors as positives (i.e. class = 1) and small reflectors as negatives (i.e. class = 0)
- Precision $= \frac{TruePositives}{TruePositives + FalsePositives}$
- Recall $= \frac{TruePositives}{TruePositives + FalseNegatives}$
- Accuracy $= \frac{TruePositives + TrueNegatives}{TruePositives + TrueNegatives + FalsePositves + FalseNegatives}$
- F1 Score $= 2 * \frac{Precision * Recall}{Precision + Recall}$

Experiments and Results

Classification

- Considering planar reflectors as positives (i.e. class = 1) and small reflectors as negatives (i.e. class = 0)
- Precision $= \frac{TruePositives}{TruePositives + FalsePositives}$
- Recall $= \frac{TruePositives}{TruePositives + FalseNegatives}$
- Accuracy $= \frac{TruePositives + TrueNegatives}{TruePositives + TrueNegatives + FalsePositves + FalseNegatives}$
- F1 Score $= 2 * \frac{Precision * Recall}{Precision + Recall}$

Dataset	Room Size	RT60	Precision	Recall	Accuracy	F1 Score
BBC Listening Room	5.6 x 5.0 x 2.9 m^3	222 ms	100 %	100 %	100 %	1.00
BBC Usability Laboratory	5.6 x 5.2 x 2.9 m^3	275 ms	83 %	100 %	89 %	0.91
Vislab	7.8 x 6.1 x 4.0 m^3	326 ms	100 %	100 %	100 %	1.00
VML	2.4 x 4.0 x 2.4 m^3	445 ms	83 %	83 %	71 %	0.83
Average	--	--	92 %	96 %	90 %	0.93

Experiments and Results

Classification

- Considering planar reflectors as positives (i.e. class = 1) and small reflectors as negatives (i.e. class = 0)
- Precision $= \frac{TruePositives}{TruePositives + FalsePositives}$
- Recall $= \frac{TruePositives}{TruePositives + FalseNegatives}$
- Accuracy $= \frac{TruePositives + TrueNegatives}{TruePositives + TrueNegatives + FalsePositves + FalseNegatives}$
- F1 Score $= 2 * \frac{Precision * Recall}{Precision + Recall}$

Dataset	Room Size	RT60	Precision	Recall	Accuracy	F1 Score
BBC Listening Room	5.6 x 5.0 x 2.9 m^3	222 ms	100 %	100 %	100 %	1.00
BBC Usability Laboratory	5.6 x 5.2 x 2.9 m^3	275 ms	83 %	100 %	89 %	0.91
Vislab	7.8 x 6.1 x 4.0 m^3	326 ms	100 %	100 %	100 %	1.00
VML	2.4 x 4.0 x 2.4 m^3	445 ms	83 %	83 %	71 %	0.83
Average	--	--	92 %	96 %	90 %	0.93

Experiments and Results

Classification

- Considering planar reflectors as positives (i.e. class = 1) and small reflectors as negatives (i.e. class = 0)
- Precision $= \frac{TruePositives}{TruePositives + FalsePositives}$
- Recall $= \frac{TruePositives}{TruePositives + FalseNegatives}$
- Accuracy $= \frac{TruePositives + TrueNegatives}{TruePositives + TrueNegatives + FalsePositves + FalseNegatives}$
- F1 Score $= 2 * \frac{Precision * Recall}{Precision + Recall}$

Dataset	Room Size	RT60	Precision	Recall	Accuracy	F1 Score
BBC Listening Room	5.6 x 5.0 x 2.9 m^3	222 ms	100 %	100 %	100 %	1.00
BBC Usability Laboratory	5.6 x 5.2 x 2.9 m^3	275 ms	83 %	100 %	89 %	0.91
Vislab	7.8 x 6.1 x 4.0 m^3	326 ms	100 %	100 %	100 %	1.00
VML	2.4 x 4.0 x 2.4 m^3	445 ms	83 %	83 %	71 %	0.83
Average	--	--	92 %	96 %	90 %	0.93

Experiments and Results

Reflector Localization

- By rendering all the reflectors as planes, the error ε is calculated as the Euclidean distance between the estimated reflector image source and its groundtruth

Experiments and Results

Reflector Localization

- By rendering all the reflectors as planes, the error ε is calculated as the Euclidean distance between the estimated reflector image source and its groundtruth
- For purposes where the listener is non-static, the importance of rendering small reflectors as cuboids instead of planes is shown by calculating the same error ε_{SARSE} by using SARSE

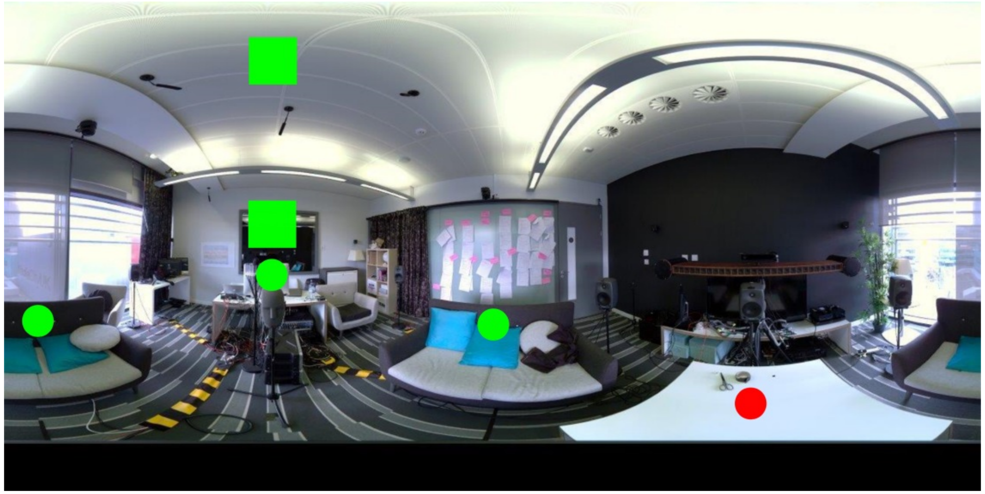
Experiments and Results

Reflector Localization

- By rendering all the reflectors as planes, the error ε is calculated as the Euclidean distance between the estimated reflector image source and its groundtruth
- For purposes where the listener is non-static, the importance of rendering small reflectors as cuboids instead of planes is shown by calculating the same error ε_{SARSE} by using SARSE

- The results reported in the table below is the average over all the image sources

Dataset	ε	ε_{SARSE}
BBC Listening Room	141 cm	**118** cm
BBC Usability Laboratory	197 cm	**150** cm
Vislab	388 cm	**178** cm
VML	**128** cm	148 cm
Average	214 cm	**149** cm



BBC | University of Salford MANCHESTER | EPSRC Pioneering research and skills | UNIVERSITY OF Southampton | UNIVERSITY OF SURREY

Conclusions

- A new method to localize acoustic reflectors given multichannel RIRs has been presented

Conclusions

- A new method to localize acoustic reflectors given multichannel RIRs has been presented

- A new binary classification algorithm has been presented to classify the acoustic reflectors by considering their size

Conclusions

- A new method to localize acoustic reflectors given multichannel RIRs has been presented
- A new binary classification algorithm has been presented to classify the acoustic reflectors by considering their size
- Results show a good accuracy provided by the localization method and the improvement given by the application of SARSE (for dynamic listening position purposes)

	Precision	Recall	Accuracy	F1 Score
Average	92 %	96 %	90 %	0.93

Conclusions

- A new method to localize acoustic reflectors given multichannel RIRs has been presented
- A new binary classification algorithm has been presented to classify the acoustic reflectors by considering their size
- Results show a good accuracy provided by the localization method and the improvement given by the application of SARSE (for dynamic listening position purposes)

	Precision	Recall	Accuracy	F1 Score
Average	92 %	96 %	90 %	0.93

- Future work may focus on extending the analysis on a greater number of rooms, finding alternative classification algorithms, and combining this method with vision-based methods.

Any Questions?

Luca Remaggi, University of Surrey, UK
l.remaggi@surrey.ac.uk