Exploiting Structural Information in Camera Aided Radar Parameter Estimation

Khurram Mazher
Sai Annaluru
Amine Mezghani
Robert W. Heath Jr.

UT Situation Aware Vehicular Engineering Systems (SAVES)
Wireless Networking and Communications Group (WNCG)
Department of Electrical and Computer Engineering
The University of Texas at Austin

This work is sponsored by
Samsung Research America

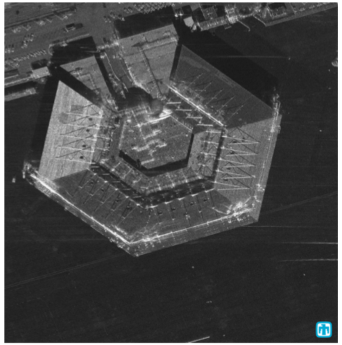
TEXAS
The University of Texas at Austin

Outline

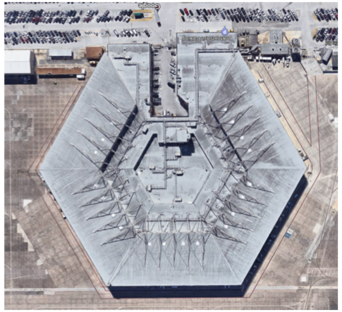
- Introduction
    - Camera radar fusion
    - Spatial spread/smoothness
    - Prior work
- Problem formulation
    - Novel regularization term
    - FFT based FMCW processing
- Results
    - Setup and Calibration
    - Experiments
- Conclusion

Introduction

Overview of Camera Radar Fusion



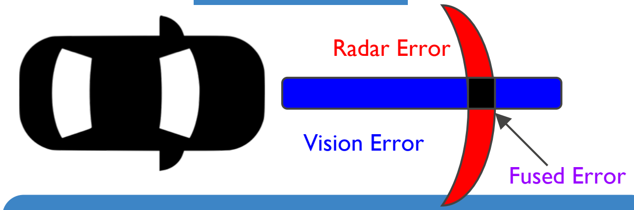
[5]



[6]

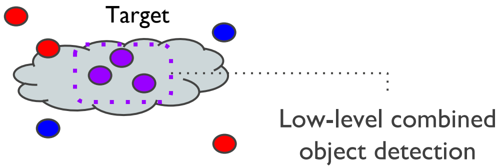
SAR and satellite image of Naval air station

Methods



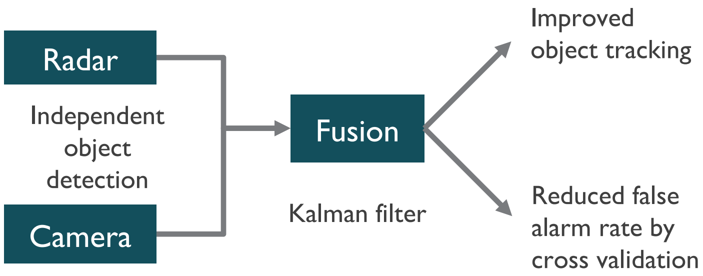
In [1], stereo vision applies contour fitting. radar applies multi-target tracking. Combined using EKF to estimate pose and motion.

Low Level Fusion



Can generate belief network of proposed object regions from both sensors and combine using an inference algorithm [2]

Post-detection fusion



Applications

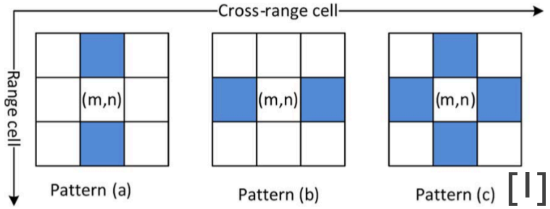
UAVs [3]

Automotive [4]

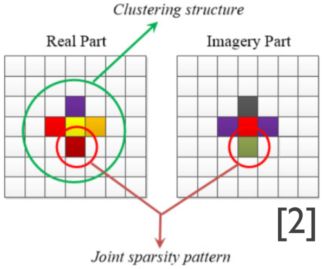
[1] S. Wu, S. Decker, P. Chang, T. Camus and J. Eledath, "Collision Sensing by Stereo Vision and Radar Sensor Fusion," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 4, pp. 606-614, Dec. 2009.
[2] B. Steux, C. Laurgeau, L. Salesse and D. Wautier, "Fade: a vehicle detection and tracking system featuring monocular color vision and radar data fusion," *Intelligent Vehicle Symposium, 2002. IEEE*, Versailles, France, 2002, pp. 632-639 vol.2.
[3] O. Meister, N. Frietsch, C. Ascher and G. F. Trommer, "Adaptive path planning for VTOL-UAVs," in IEEE Aerospace and Electronic Systems Magazine, vol. 24, no. 7, pp. 36-41, July 2009.
[4] X. Wang, L. Xu, H. Sun, J. Xin and N. Zheng, "On-Road Vehicle Detection and Tracking Using MMW Radar and Monovision Fusion," in IEEE Transactions on ITS, vol. 17, no. 7, pp. 2075-2084, July 2016.
[5] https://www.sandia.gov/radar/imagery/index.html and [6] Google maps satellite image of same location

Clustering: structure in addition to sparsity

Radar



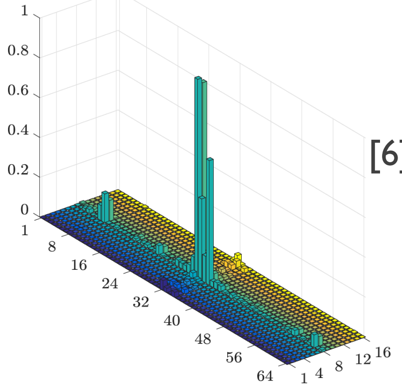
[1]



[2]

Clustering/smoothness [1,2,3] exploited in addition to sparsity in AoA, AoD etc

Communication



[6]

Clustering in AoA, AoD & delay has been observed in prior work [7,8]

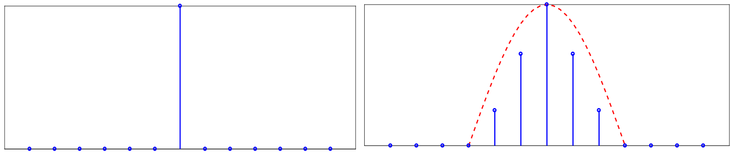
Other applications

Audio signals [3]

Music [4]

EEG [5]

Off-grid parameter estimation (leakage)



Block sparse: $[\underbrace{x_1\ x_2\ x_3}\ \ldots\ \underbrace{x_{i-1}\ x_i\ x_{i+1}}\ \ldots\ x_n\ ]^T$

Non-zeros cluster/block

Instead of being sporadically sparse, *nature* is smooth with non-zero entries occurring in clusters. This can be exploited for estimation problems.

[1] L. Wang, L. Zhao, G. Bi, C. Wan and L. Yang, "Enhanced ISAR Imaging by Exploiting the Continuity of the Target Scene," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5736-5750, Sept. 2014.
[2] X. Wang, G. Li, Y. Liu and M. G. Amin, "Enhanced 1-Bit Radar Imaging by Exploiting Two-Level Block Sparsity," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 2, pp. 1131-1141, Feb. 2019.
[3] H. Duan, L. Zhang, J. Fang, L. Huang and H. Li, "Pattern-Coupled Sparse Bayesian Learning for Inverse Synthetic Aperture Radar Imaging," in *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1995-1999, Nov. 2015.
[4] L. Yu, H. Sun, J.P. Barbot, and G.Zheng, "Bayesian compressive sensing for cluster structured sparse signals," in *Signal processing* 92, no. 1 (2012): 259-269.
[5] Z. Zhang and B. D. Rao, "Sparse Signal Recovery With Temporally Correlated Source Vectors Using Sparse Bayesian Learning," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912-926, Sept. 2011.
[6] A. Ali, N. González-Prelcic and R. W. Heath, "Millimeter Wave Beam-Selection Using Out-of-Band Spatial Information," in *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1038-1052, Feb. 2018.
[7] P. Schniter, "A Message-Passing Receiver for BICM-OFDM Over Unknown Clustered-Sparse Channels," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 8, pp. 1462-1474, Dec. 2011.
[8] P. Wang, M. Pajovic, P. V. Orlik, T. Koike-Akino, K. J. Kim and J. Fang, "Sparse channel estimation in millimeter wave communications: Exploiting joint AoD-AoA angular spread," *2017 IEEE International Conference on Communications (ICC)*, 2017.

Prior work

LASSO

$$||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2 + \lambda||\mathbf{x}||_1$$

Sparsity inducing ℓ_1-norm minimization solution (LASSO) – picks out the largest entries of $\mathbf{x}$ depending on λ

PC-SBL , $\beta = 0$

Sparse Bayesian learning with independent prior [1]. A Bayesian way of forming a sparse solution

PC-SBL , $\beta = 1$

Pattern coupled sparse Bayesian learning with correlated prior [1]. Continuity encouraged by sharing hyperparameters of underlying Gaussian prior between neighboring entries

Coupled ℓ_1-norm

Proposed solution: Novel regularization of linear model that combines sparsity and continuity. We use the term "Coupled ℓ_1-norm" for this.

[1] J. Fang, Y. Shen, H. Li and P. Wang, "Pattern-Coupled Sparse Bayesian Learning for Recovery of Block-Sparse Signals," in *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 360-372, Jan.15, 2015.

Problem formulation

Soft-sparsity regularization

Captures many formulations in communications, radar, navigation

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$$

- Linear system solution
- Sparse representation
- Compressed sensing measurement

$$J(\mathbf{x}) = \underbrace{||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2}_{\text{Model constraint}} + \lambda \underbrace{\sqrt{|\mathbf{x}|}^T \mathbf{R}^{-1} \sqrt{|\mathbf{x}|}}_{\text{"Soft-sparsity" constraint}}$$

Element-wise "coupled ℓ_1" norm

$\mathbf{R}$ based on structural information or side information

For $\mathbf{R} = \mathbf{I}$, the formulation simplifies to ℓ_1-norm regularization

$$J(\mathbf{x}) = ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2^2 + \lambda \sqrt{|\mathbf{x}|}^T \sqrt{|\mathbf{x}|}$$

General $\mathbf{R}$: Introducing off-diagonal terms in $\mathbf{R}$ can be thought of as a generalization of ℓ_1-norm where R signifies the correlation between the values of amplitude of $\mathbf{x}$

Extension to 2D

Beamspace channel model

$$\mathbf{Y} = \mathbf{A}_{\mathrm{L}}\mathbf{X}\mathbf{A}_{\mathrm{R}} + \mathbf{N}$$

Range DoA radar processing

Left measurement\dictionary matrix

Right measurement\dictionary matrix

$$J(\mathbf{X}) = \underbrace{||\mathbf{Y} - \mathbf{A}_{\mathrm{L}}\mathbf{X}\mathbf{A}_{\mathrm{R}}||_{\mathrm{F}}^{2}}_{\text{Model constraint}} + \lambda\, \underbrace{\mathrm{Tr}\left(\mathbf{R}_{\mathrm{L}}^{-1}\sqrt{|\mathbf{X}|^{T}}\mathbf{R}_{\mathrm{R}}^{-1}\sqrt{|\mathbf{X}|}\right)}_{\text{“Soft-sparsity” constraint}}$$

Element-wise

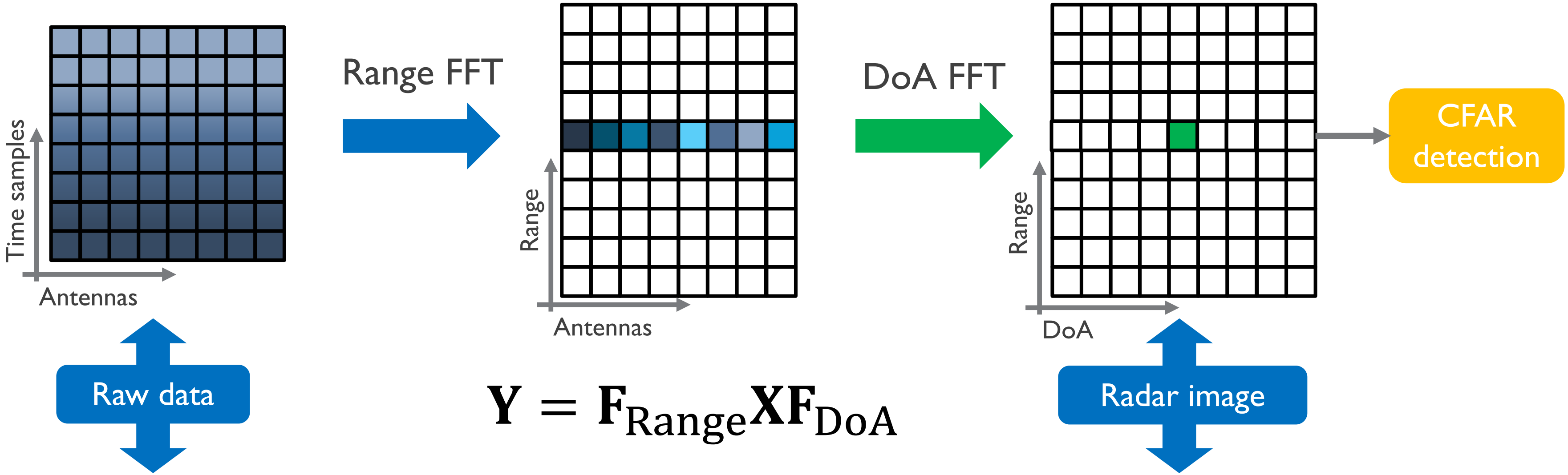
This model captures the FFT based range-DoA processing on a FMCW radar

Solution

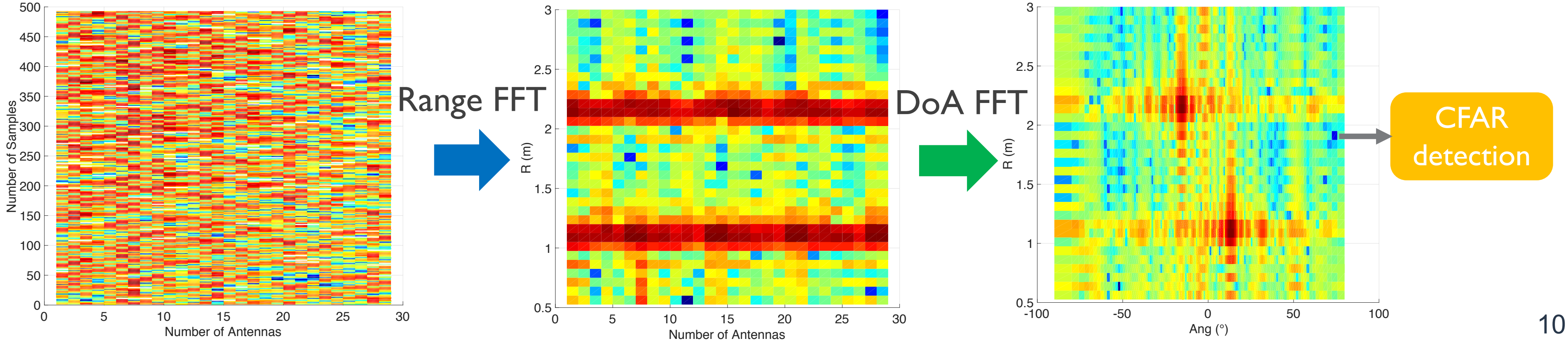
$$\nabla J(\mathbf{X}) = -\mathbf{A}_{\mathrm{L}}^{\mathrm{T}}\mathbf{Y}^{\mathrm{c}}\mathbf{A}_{\mathrm{R}}^{\mathrm{T}} + \mathbf{A}_{\mathrm{L}}^{\mathrm{T}}\mathbf{A}_{\mathrm{L}}^{\mathrm{c}}\mathbf{X}^{\mathrm{c}}\mathbf{A}_{\mathrm{R}}^{\mathrm{c}}\mathbf{A}_{\mathrm{R}}^{\mathrm{T}} \quad +$$
$$\frac{\lambda}{2}\left(\mathbf{R}_{\mathrm{R}}^{-1}(|\mathbf{X}|)^{\frac{1}{2}}\mathbf{R}_{\mathrm{L}}^{-1} + (\mathbf{R}_{\mathrm{R}}^{-1})^{\mathrm{T}}(|\mathbf{X}|)^{\frac{1}{2}}(\mathbf{R}_{\mathrm{L}}^{-1})^{\mathrm{T}}\right) \odot \left(1./(|\mathbf{X}|)^{\frac{1}{2}}\right) \odot e^{\angle(\mathbf{X}^{\mathrm{c}})}$$

Gradient descent based iterative solution to the above optimization

FFT based FMCW processing



$$\mathbf{Y} = \mathbf{F}_{\mathrm{Range}}\mathbf{X}\mathbf{F}_{\mathrm{DoA}}$$



Matrix R

$$\mathbf{R}(a)_{N\times N} = \begin{bmatrix} 1 & a & 0 & \dots & 0 \\ a & 1 & a & \dots & 0 \\ 0 & a & 1 & a & \dots \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & a & 1 \end{bmatrix}_{N\times N}$$

$$\mathbf{R} = \begin{bmatrix} 1 & 0.4 & 0.1 & \dots & 0 & 0 & 0 \\ 0.4 & 1 & 0.4 & 0.1 & \dots & 0 & 0 \\ 0.1 & 0.4 & 1 & 0.4 & 0.1 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & \dots & 0.1 & 0.4 & 1 \end{bmatrix}_{N\times N} \quad \mathbf{R}^{-1} \approx \begin{bmatrix} 1 & -0.38 & 0.05 & \dots & 0 & 0 & 0 \\ -0.38 & 1 & -0.38 & 0.05 & \dots & 0 & 0 \\ 0.05 & -0.38 & 1 & -0.38 & 0.05 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \ddots & \ddots & 0.05 & -0.38 & 1 \end{bmatrix}_{N\times N}$$

Structural information approximated by a *Toeplitz* matrix as a heuristic

Encourage recovery in this area

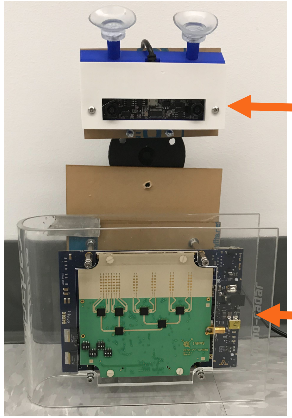
$$\mathbf{R} = \begin{bmatrix} 10 & 4 & 0 & \dots & 0 & 0 \\ 4 & 10 & 4 & 0 & \dots & 0 \\ 0 & 4 & 10 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & 0 & \dots & \dots & 0 & 1 \end{bmatrix}_{N\times N}$$

To exploit prior knowledge, one could give unequal weight on the diagonal

The choice of $\mathbf{R}$ depends on structural knowledge about the problem and any side information. Our current choice is based on heuristics and experimentation.

Experiments

Experimental setup



Stereo camera

INRAS RadarBook (FMCW)

Parameter	Value
Center Frequency f_c	76.5 GHz
Sweep Bandwidth	1 GHz
# of samples N	504
N_{TX}	4 (7λ /2 Spacing)
N_{RX}	8 (λ /2 Spacing)
# elements in virtual aperture	29 (λ /2 Spacing)
Camera Resolution	1080p

Radar only setup

- Capture radar data
- **R** matrix formulation based on heuristic

Radar + camera setup

- Calibration
- Capture camera and radar data
- Camera object detection (YOLO [1])
- **R** matrix formulation based on detected object and calibration

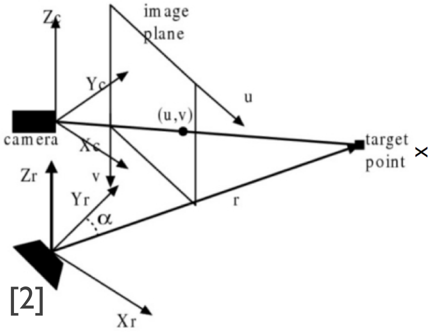
Coupled ℓ_1 solver

Other/better "recipes" for construction of **R** possible. Ongoing work.

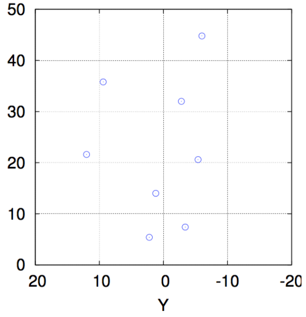
[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 779–788.

Calibration

Extrinsic calibration [1] (B/w one camera and radar)



[2]



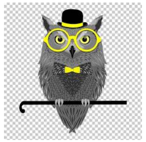
[1]

$$\begin{bmatrix} x_r \\ y_r \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix}$$

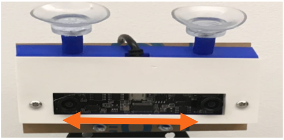
Homography: A 2D projective transformation that maps the camera pixels to radar pixels (at least spatially – degenerate in range dimension)

Calibration helps transform the camera co-ordinate system to the radar co-ordinate system

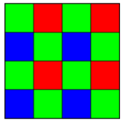
Stereo calibration [3] and depth estimation (B/w two cameras)



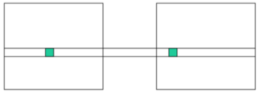
Stereo



b = Baseline



ps = Pixel size



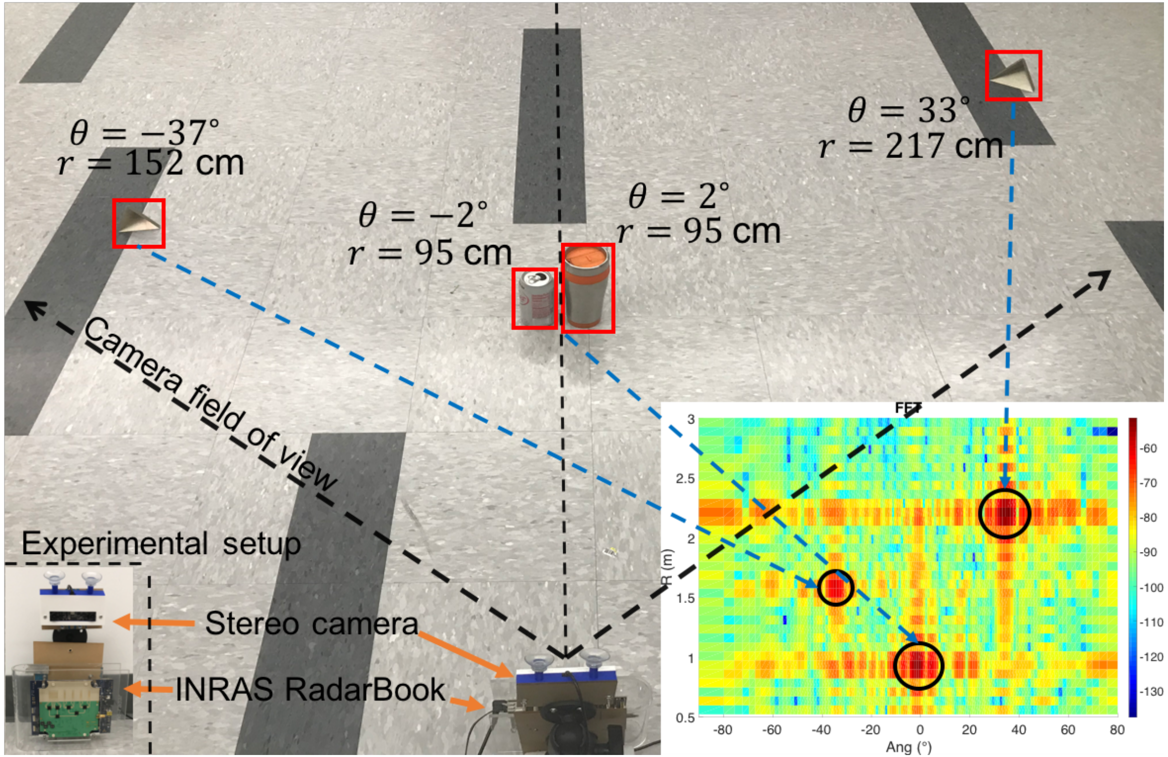
d = Disparity value

$$D = \frac{f * b}{d * ps}$$

Translation and rotational matrices b/w the two cameras

[1] J. Oh, K. Kim, M. Park and S. Kim, "A Comparative Study on Camera-Radar Calibration Methods," *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Singapore, 2018, pp. 1057-1062.
[2] [1] G. Bradski, "The OpenCV Library,"Dr. Dobb's Journal of Software Tools, Internet: https://www.vision-systems.com/content/dam/VSD/NextGen/5-3D-2.pdf, 2000
[3] Stereo camera calibrator https://www.mathworks.com/help/vision/ref/stereocameracalibrator-app.html

2D scenario: Range and Spatial DoA

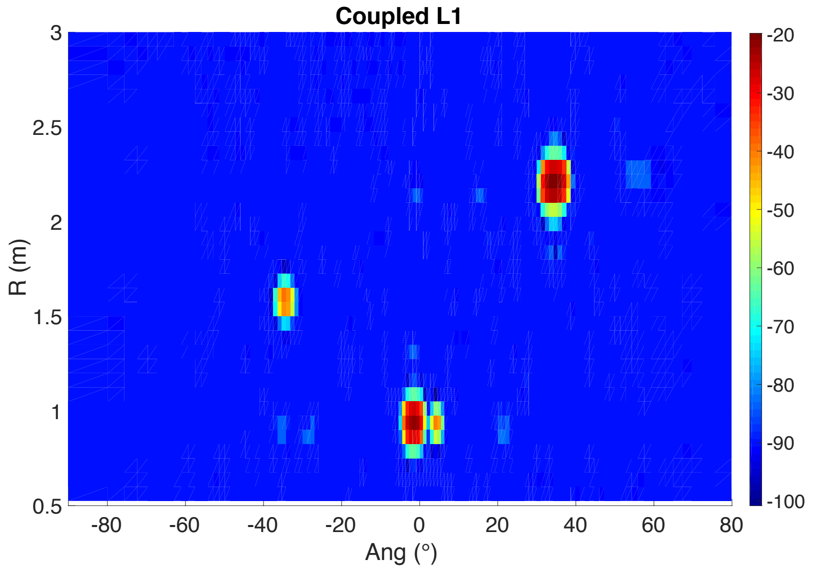


No 'ground truth' available here

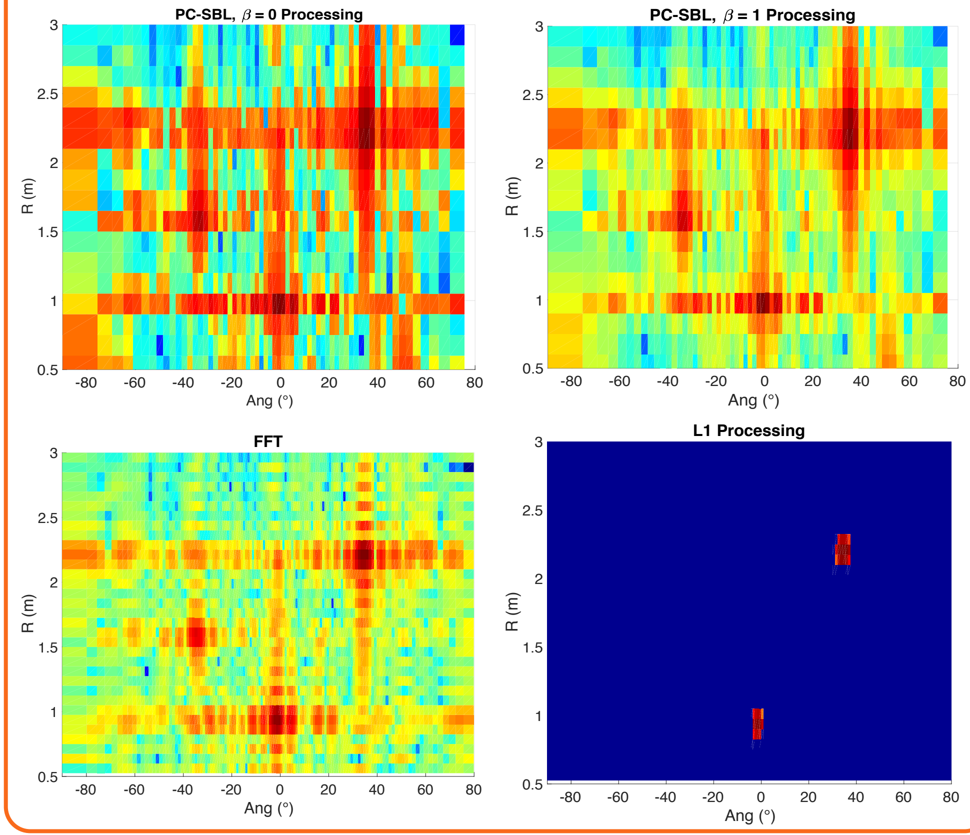
For now the 'ground truth' is taken as the FFT based processing

Radar only setting

Proposed solution

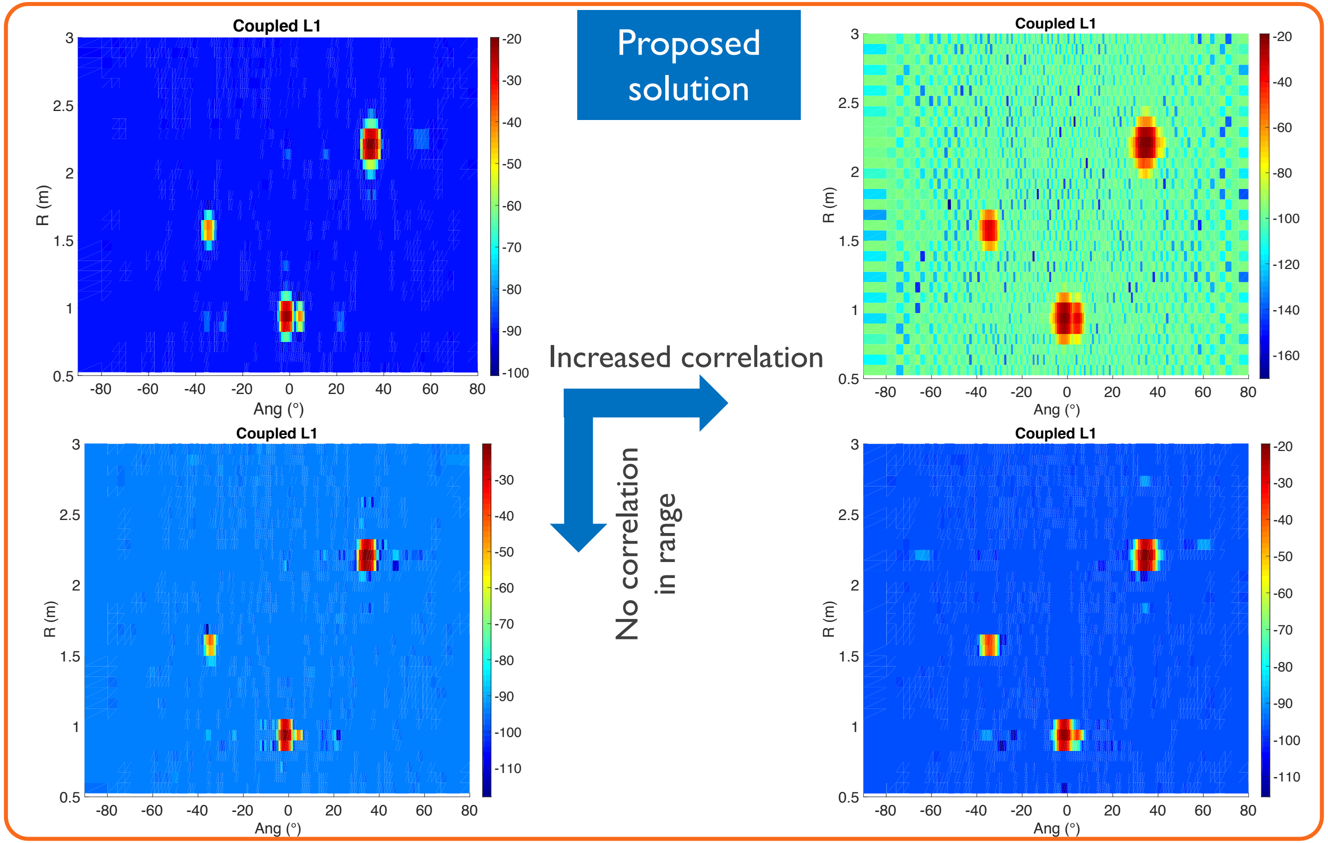


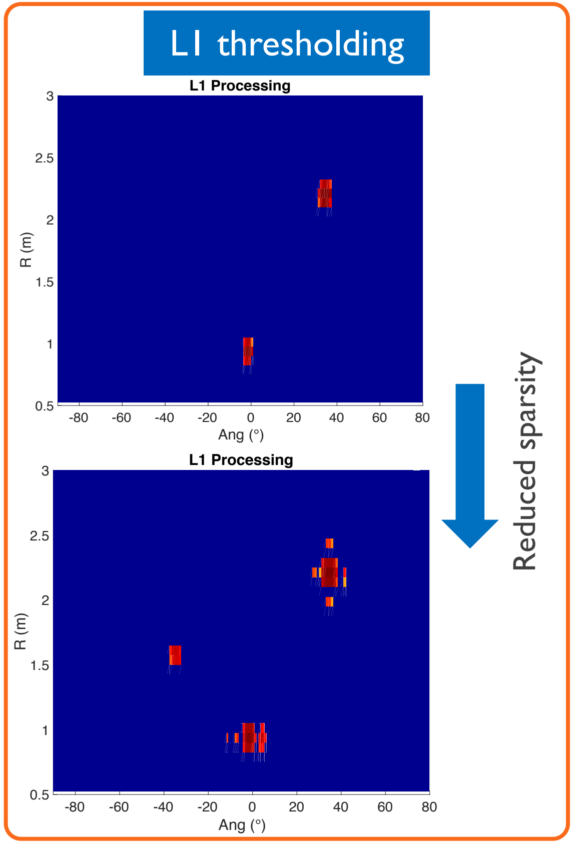
Comparison algorithms



Targets can be observed clearly separated from the background clutter. Additionally, the two targets placed close-by can be resolved.

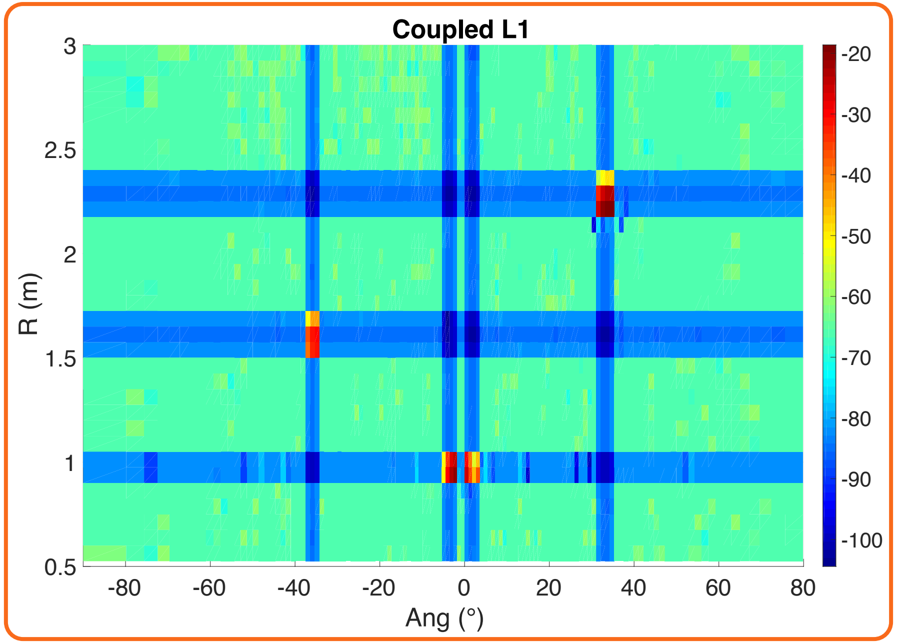
Effect of R matrix





The correlation assumed amongst the values of $\mathbf{x}$ controls the spatial spread of the radar image. The matrix $\mathbf{R}$ needs to be chosen carefully.

Radar + stereo camera setting



Better separation in both DoA and range. Depends on the range and DoA of potential target reported by camera. Minimal side-lobes.

Object detection in image (black box e.g. YOLO [1])

Map the pixels of the detected objects to spatial DoA using extrinsic calibration data

Compute distance to objects using calibration between the 2 cameras

$$b \times \begin{bmatrix} 1 & a & 0 & \dots & 0 \\ a & 1 & a & \ddots & \ddots \\ 0 & a & 1 & a & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & a & 1 \end{bmatrix}$$

Place Toeplitz sub-matrices in the bins of $\mathbf{R}_{\mathrm{R}}$ corresponding to the DoA of the objects

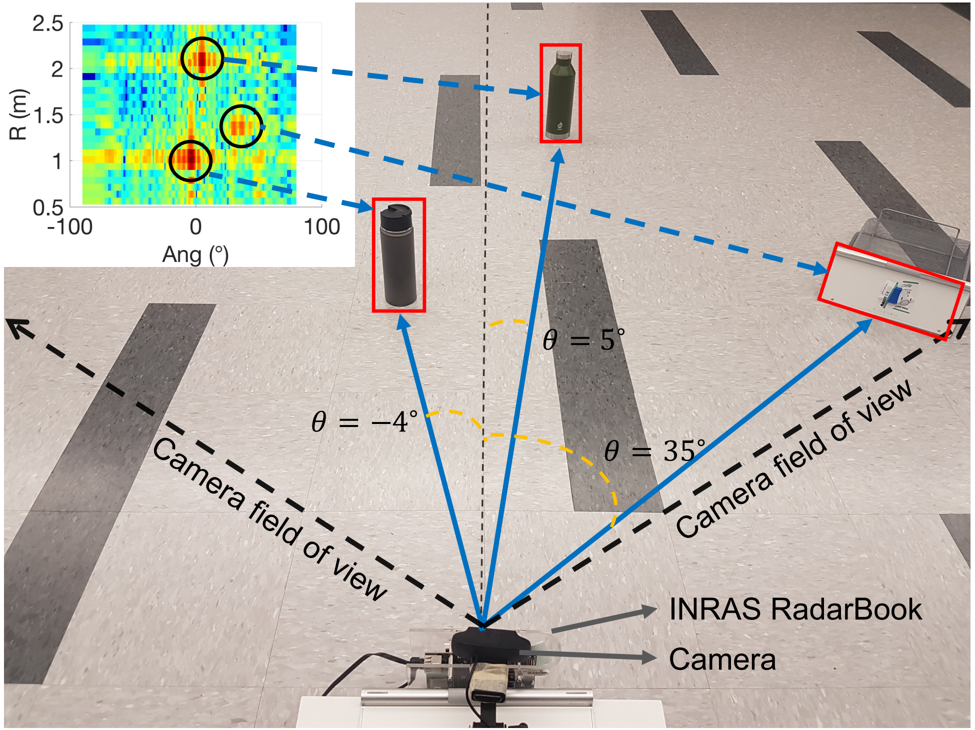
Place Toeplitz sub-matrices at appropriate bins of $\mathbf{R}_{\mathrm{L}}$

Recipie for $\mathbf{R}_{\mathrm{R}}$

Recipie for $\mathbf{R}_{\mathrm{L}}$

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016, pp. 779–788.

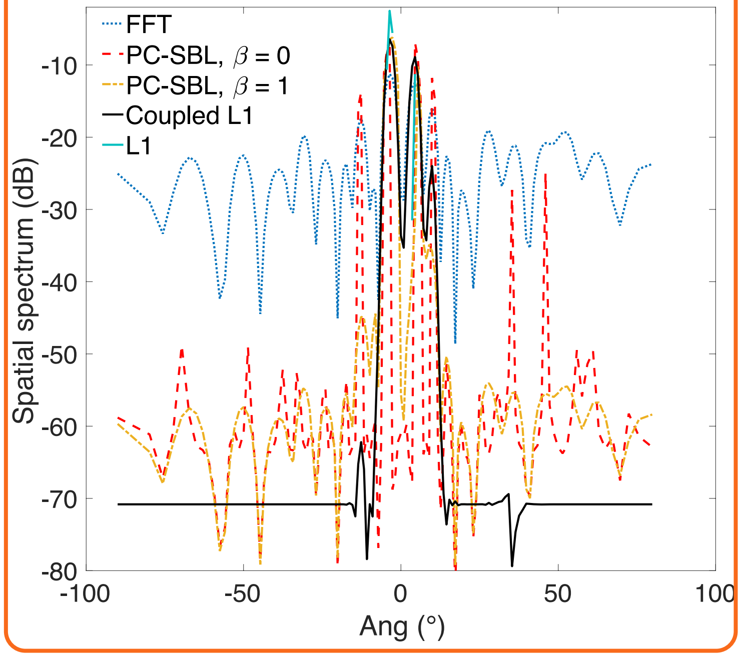
1D scenario: Spatial DoA



1D spatial data obtained by summing along the range dimension. All targets collapse onto the DoA dimension.

1D spatial results

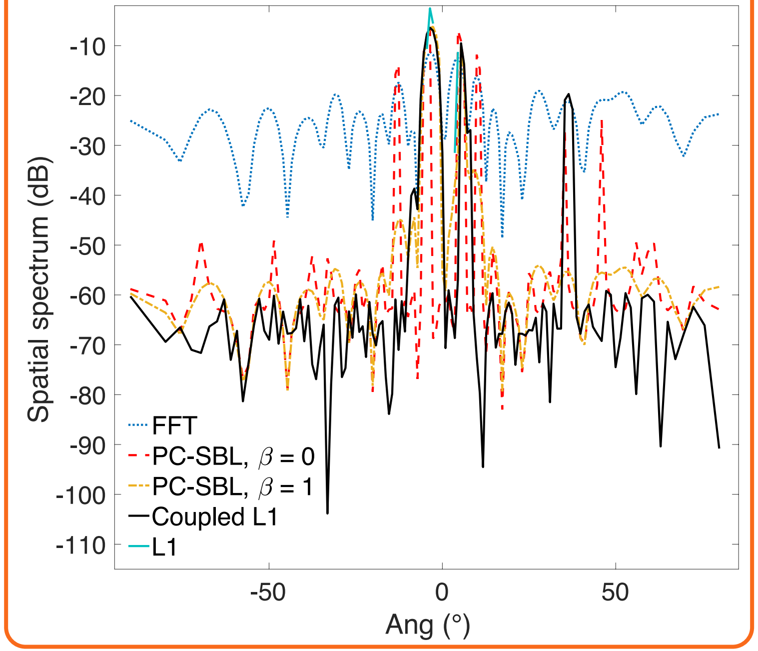
Radar only



Coherent with a smooth version of L1 solution

Solution biased in favor of the side information from camera module

Radar + camera



The proposed solution achieves a good balance between the side information from the secondary sensor, prior assumed structural information and the underlying measurements through the matrix $\mathbf{R}$.

Conclusion

Conclusion

Novel regularization	$\sqrt{|\mathbf{x}|}^T \mathbf{R}^{-1} \sqrt{|\mathbf{x}|}$ allows for easy incorporation of *prior* structural information and *side* information through the matrix $\mathbf{R}$
Design of $\mathbf{R}$	Heuristic design based on *Toeplitz* matrices. Recipe for design based on camera.
Experiments	Experimental results illustrate improved clutter separation, increased resolution and greater sensitivity towards side information
Generic	The proposed formulation applies to linear problems in communication & navigation.

Ongoing work …

Optimize $\mathbf{R}$	Better ways to design $\mathbf{R}$ …?
Algorithmic improvements	Current implementation is based on gradient descent. We are exploring alternatives based on VAMP