Research @ Visual Analytics

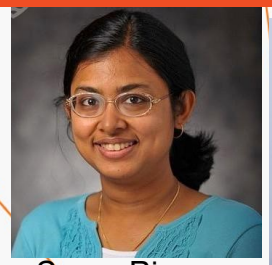
Presenter: Soma Biswas

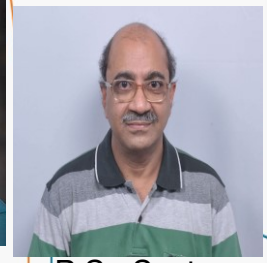
Department of Electrical Engineering

Email: somabiswas@iisc.ac.in

AG Ramakrishnan

Venu Govindu


Soma Biswas


P.S. Sastry

Chandra Sekhar Seelamantula

Kunal Chaudhury

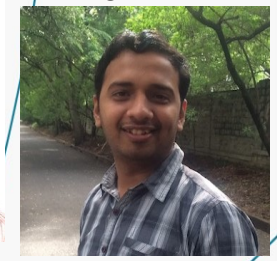
Prasanta Kumar Ghosh

Muthuvel Arigovindan

Anirban Chakraborty

R. Venkatesh Babu

Phaneendra K. Yalavarthy


Rajiv Soundararajan

Chiranjib Bhattacharyya

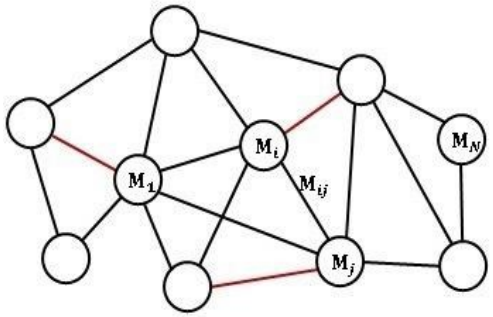
Vijay Natarajan

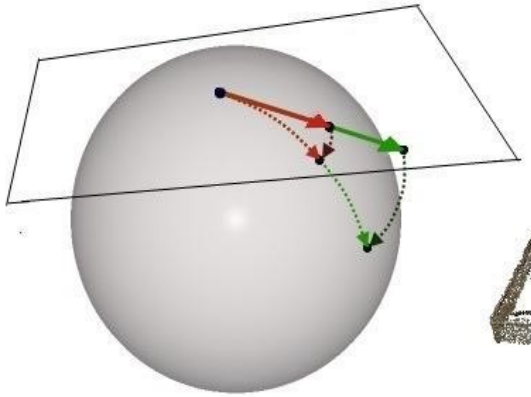
Ambedkar Dukkipati

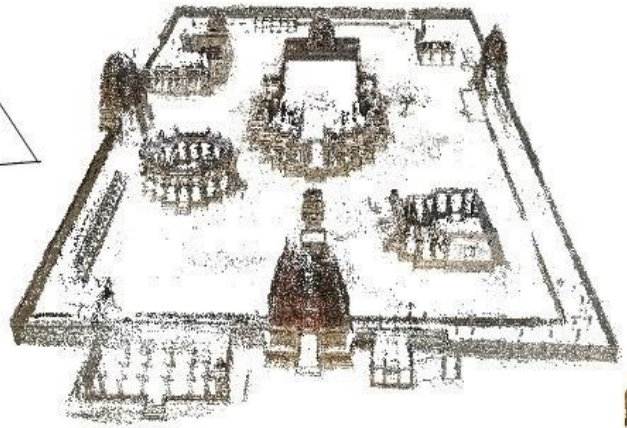
Geometry in Computer Vision

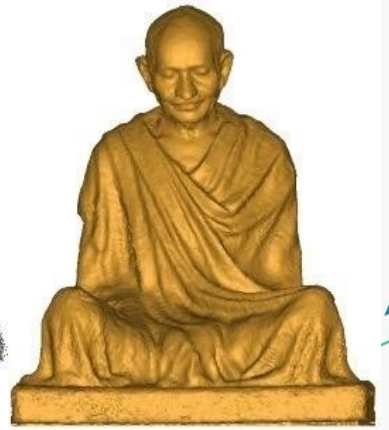
- Estimation on geometric manifolds
- Large-scale 3D reconstruction
- High quality 3D scanning using depth cameras
- Geometry at the intersection of SLAM and SfM.

- New Framework: r**otation averaging**
- efficient, scalable, robust, and accurate
- better than state of the art

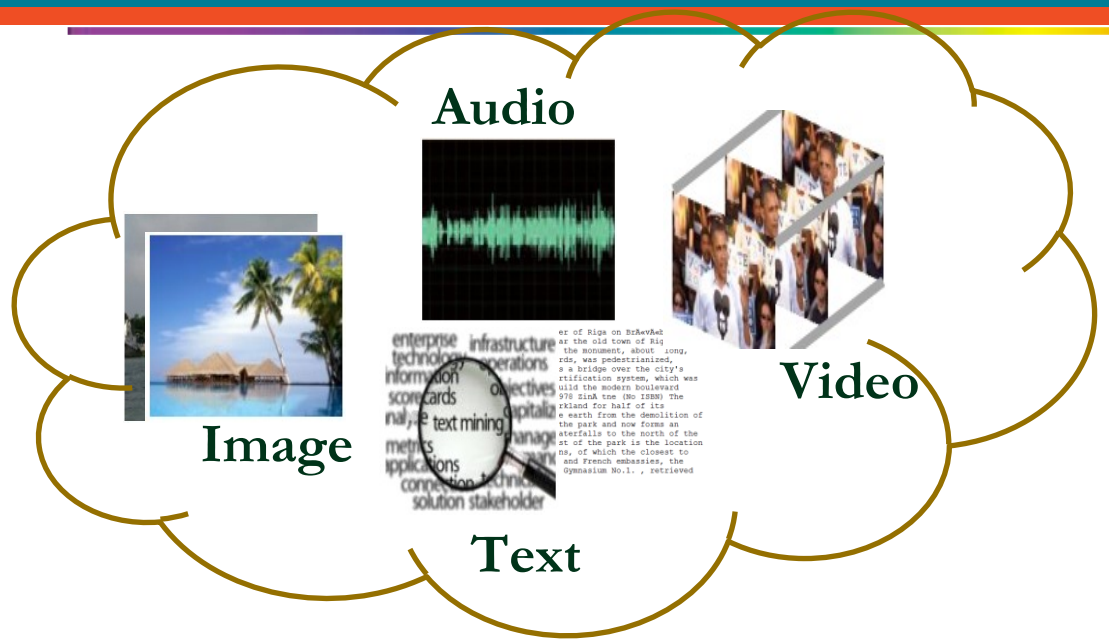


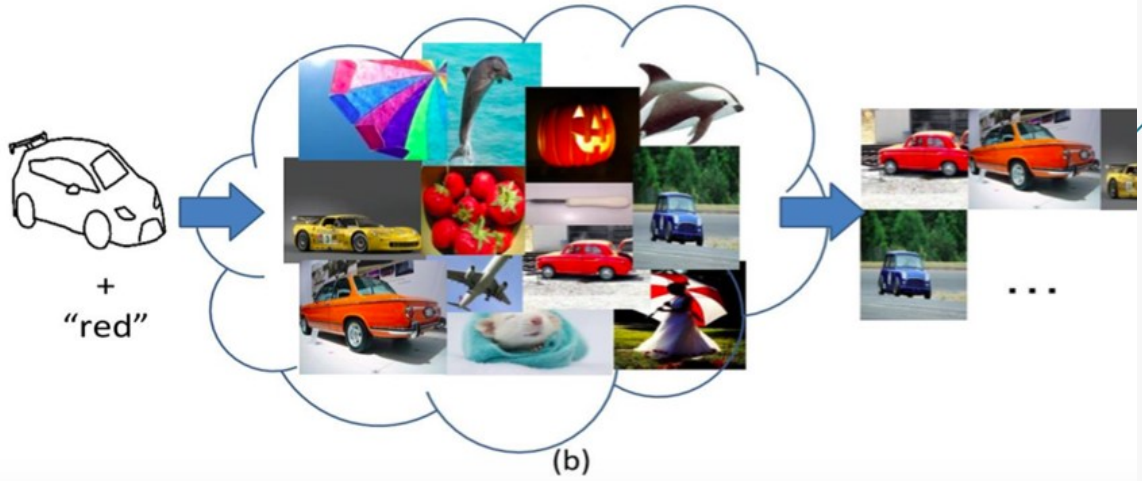




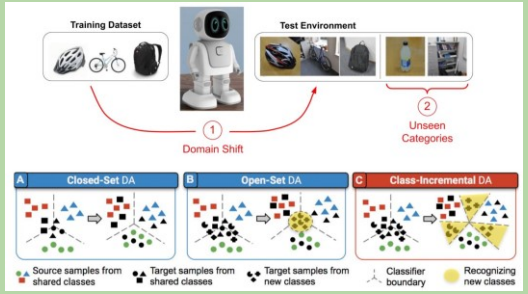


Cross-Modal Matching in Computer Vision

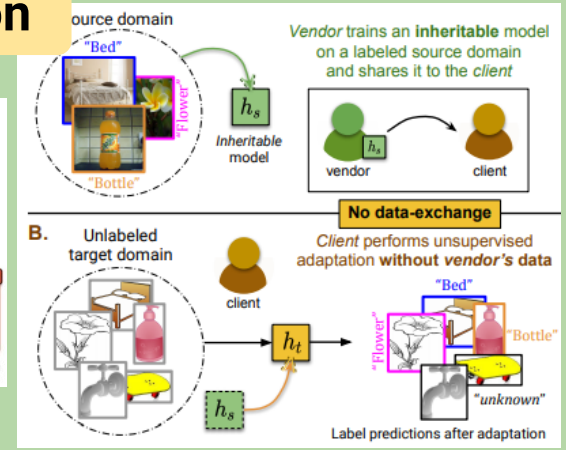




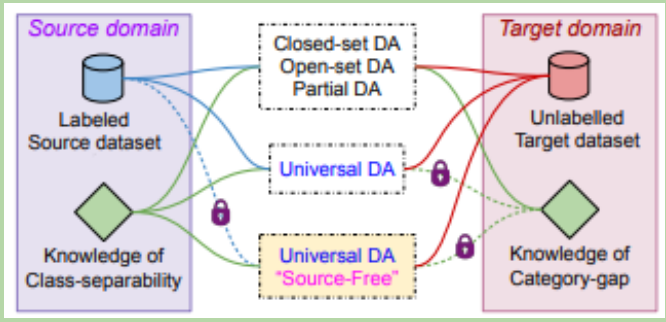
Domain Adaptation



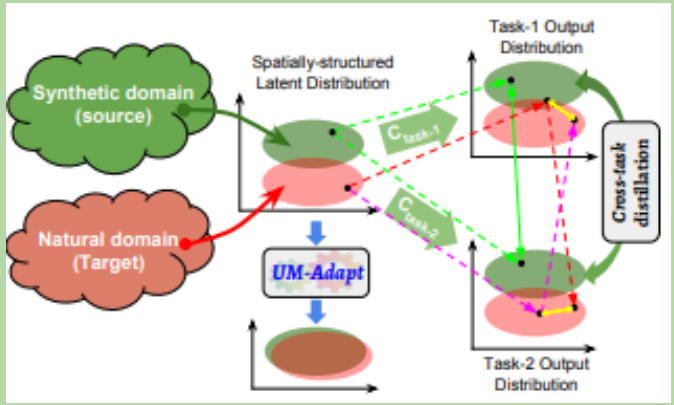
"Class-Incremental Domain Adaptation", Kundu *et al.*, ECCV 2020



"Towards Inheritable Models for Open-set Domain Adaptation", Kundu *et al.*, CVPR 2020

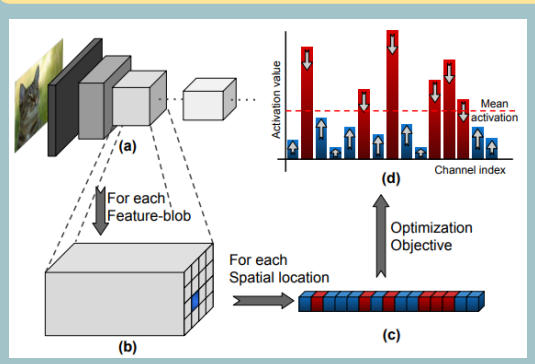


"Universal Source-Free Domain Adaptation", Kundu *et al.*, CVPR 2020

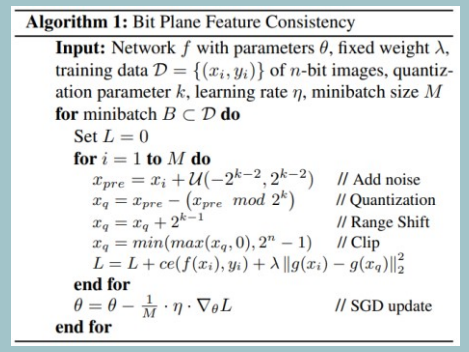


"UM-Adapt: Unsupervised Multi-task Adaptation using Adversarial Cross-task Distillation", Kundu *et al.*, ICCV 2019

Adversarial Machine Learning



"Feature Disruptive attack", Adhitya *et al.*, ICCV 2019

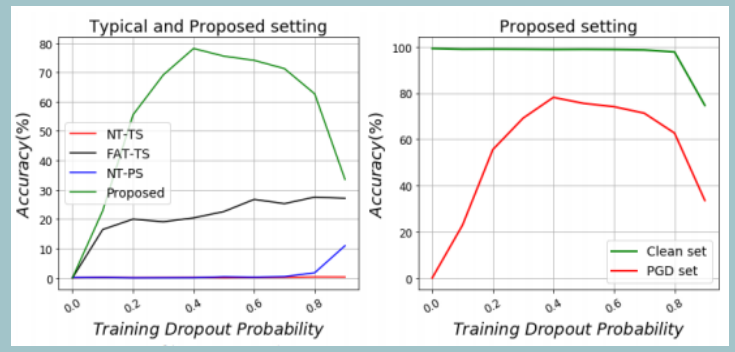


Algorithm 1: Bit Plane Feature Consistency

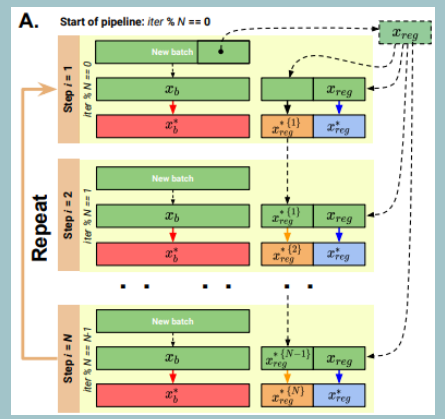
Input: Network f with parameters θ, fixed weight λ, training data $\mathcal{D} = \{(x_i, y_i)\}$ of n-bit images, quantization parameter k, learning rate η, minibatch size M

for minibatch $B \subset \mathcal{D}$ **do**
  Set $L = 0$
  for $i = 1$ **to** M **do**
    $x_{pre} = x_i + \mathcal{U}(-2^{k-2}, 2^{k-2})$ // Add noise
    $x_q = x_{pre} - (x_{pre} \mod 2^k)$ // Quantization
    $x_q = x_q + 2^{k-1}$ // Range Shift
    $x_q = min(max(x_q, 0), 2^n - 1)$ // Clip
    $L = L + ce(f(x_i), y_i) + \lambda \|g(x_i) - g(x_q)\|_2^2$
  end for
  $\theta = \theta - \frac{1}{M} \cdot \eta \cdot \nabla_\theta L$ // SGD update
end for

"Towards Achieving Adversarial Robustness by Enforcing Feature Consistency Across Bit Planes", Sravanti *et al.*, CVPR 2020

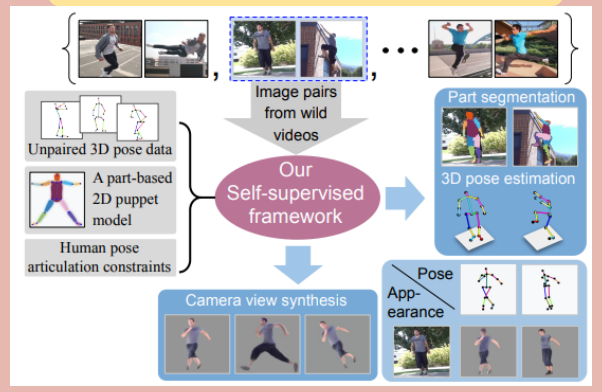


"Single-step Adversarial training with Dropout Scheduling", Vivek *et al.*, CVPR 2020

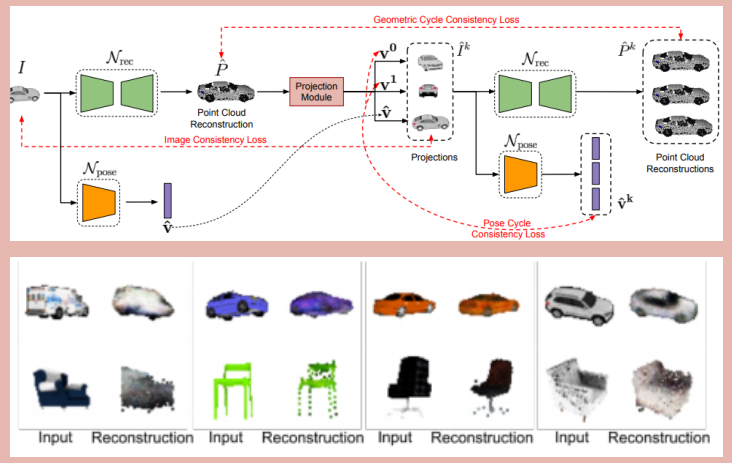


"Plug-And-Pipeline: Efficient Regularization for Single-Step Adversarial Training", Vivek *et al.*, CVPRW 2020

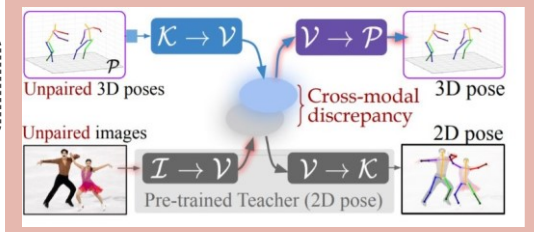
3D Computer Vision

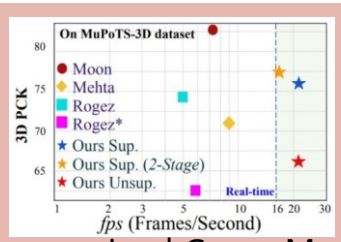


"Self-Supervised 3D Human Pose Estimation Via Part Guided Novel Image Synthesis", Kundu *et al.*, CVPR 2020

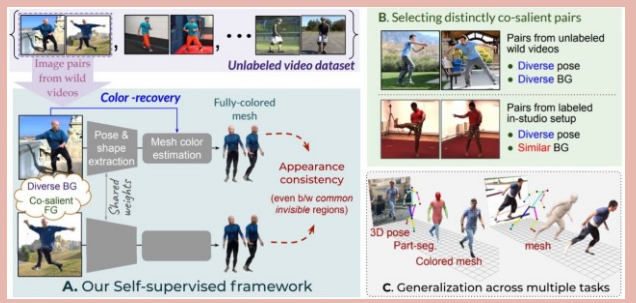


"From Image Collections to Point Clouds with Self-supervised Shape and Pose Networks", Navaneet *et al.*, CVPR 2020



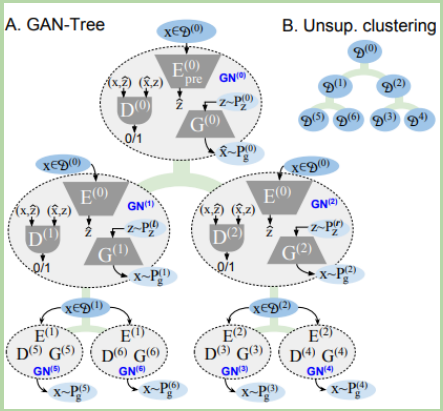


"Unsupervised Cross-Modal Alignment For Multi-Person 3D Pose Estimation", Kundu *et al.*, ECCV 2020



"Appearance Consensus Driven Self-Supervised Human Mesh Recovery", Kundu *et al.*, ECCV 2020

Generative Adversarial Networks

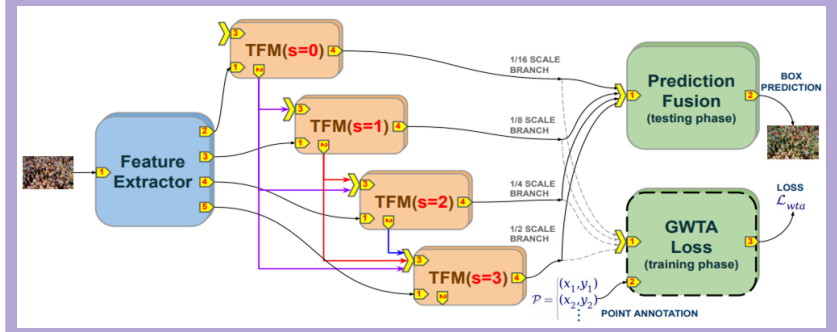


"GAN-Tree: An Incrementally Learned Hierarchical Generative Framework for Multi-modal Data Distributions", Kundu *et al.*, ICCV 2019

Computational Photography

"Towards Practical and Efficient High-Resolution HDR Deghosting with CNN", Prabhakar *et al.*, ECCV 2020

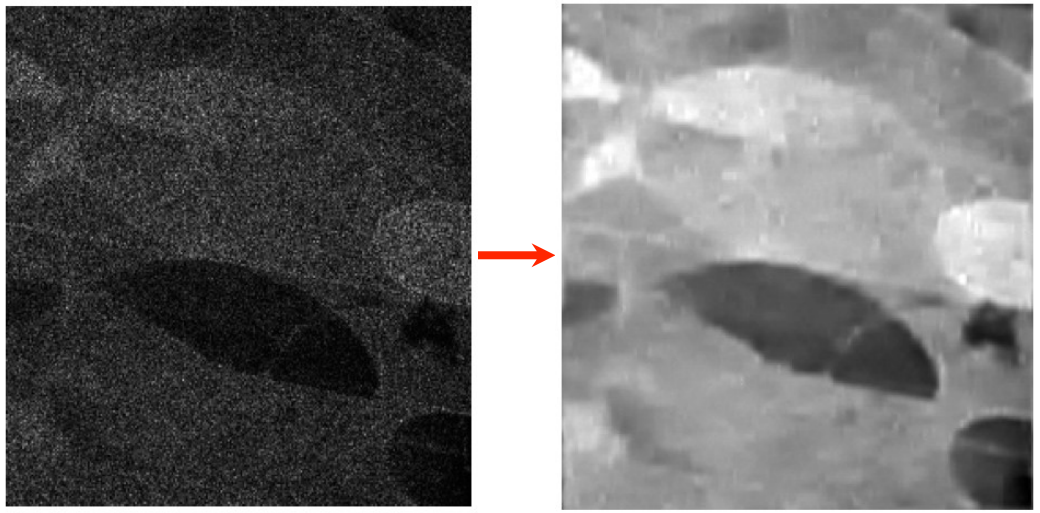
Crowd Counting

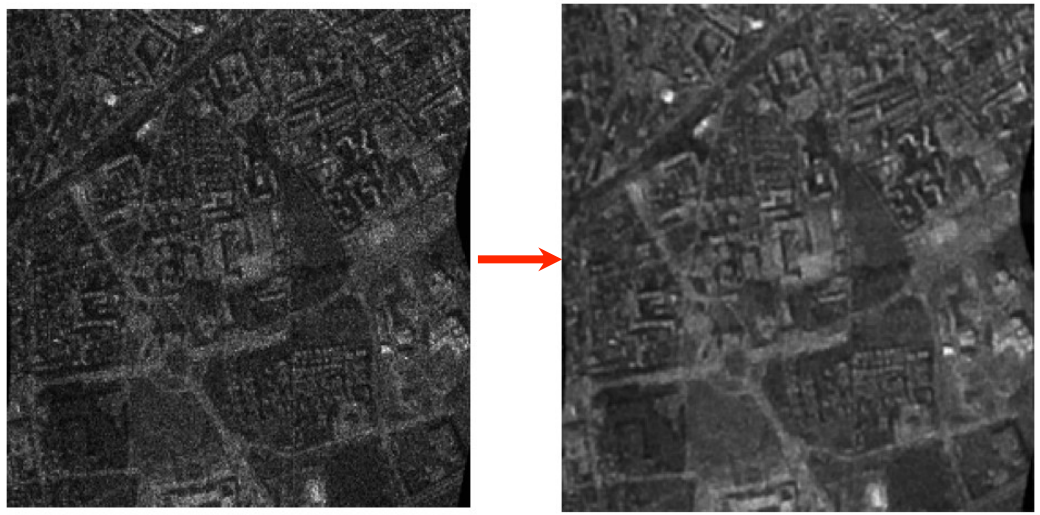


"Locate, Size and Count: Accurately Resolving People in Dense Crowds via Detection", Deepak *et al.*, TPAMI 2020

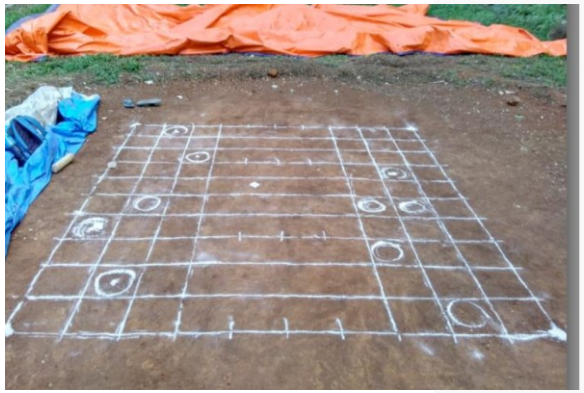
Image Restoration & Super-resolution Imaging

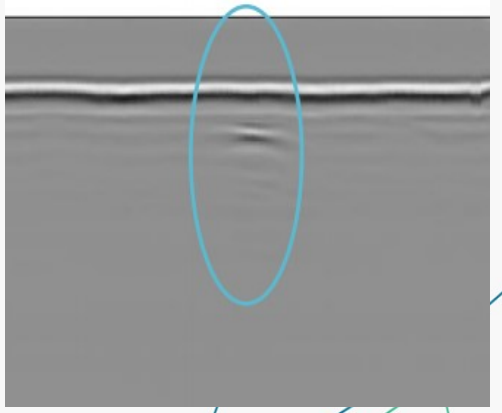
Restoration of satellite images





Ground penetrating radar
EE & Civil Eng., IISc

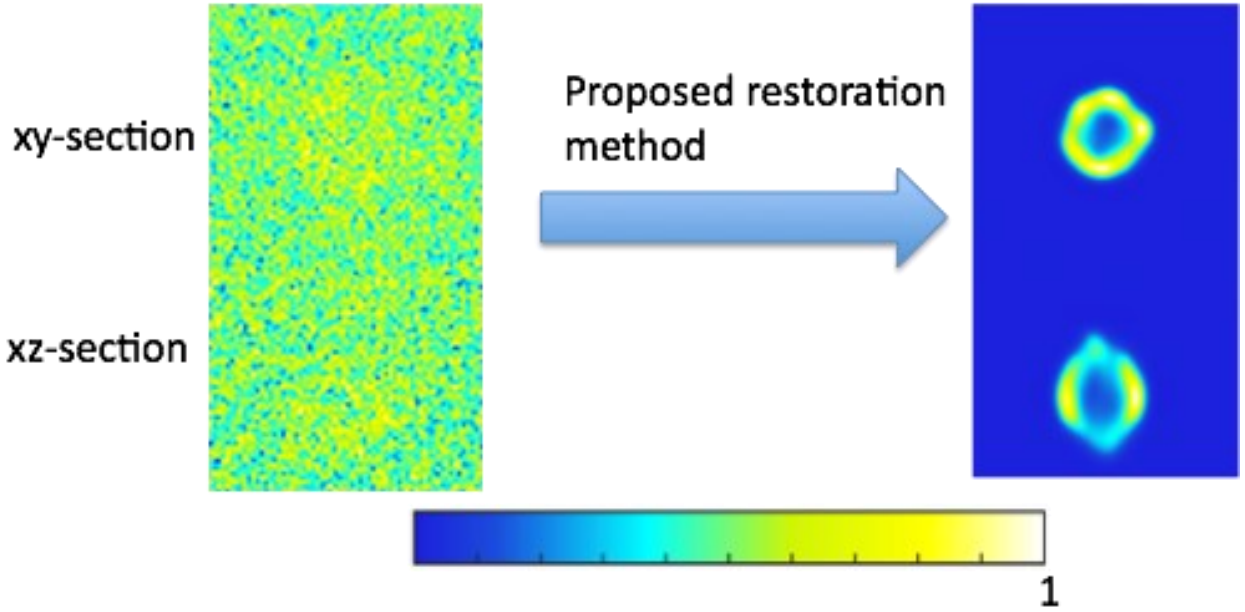


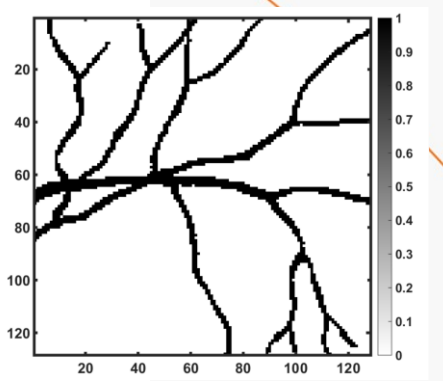


- Landmine detection
- Low-cost GPR, Low-frequency GPR? (800 MHz, instead of 1.5 GHz?)

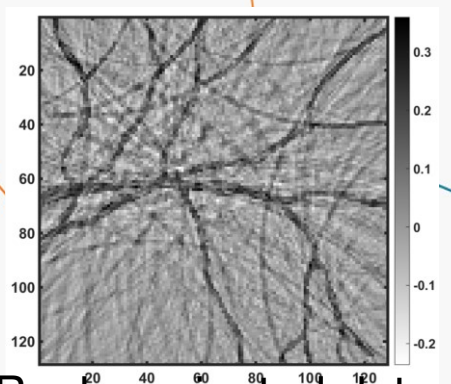
Image Restoration in Optical Imaging

- Image restoration, structurally adaptive regularization
- Fluorescence microscopy & photoacoustic imaging
- Redesigning confocal microscopy system for reduced photodamage
- Optical systems modeling and optical telescopy

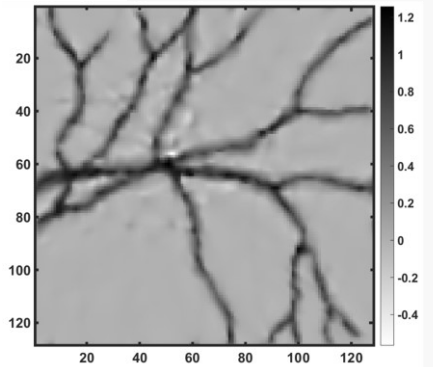




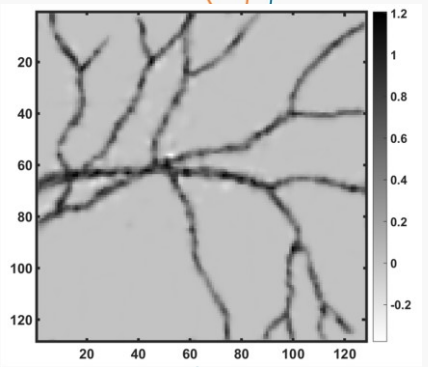
Phantom model



Back-projected data from 32 transducers



Second-order TV (most widely used)

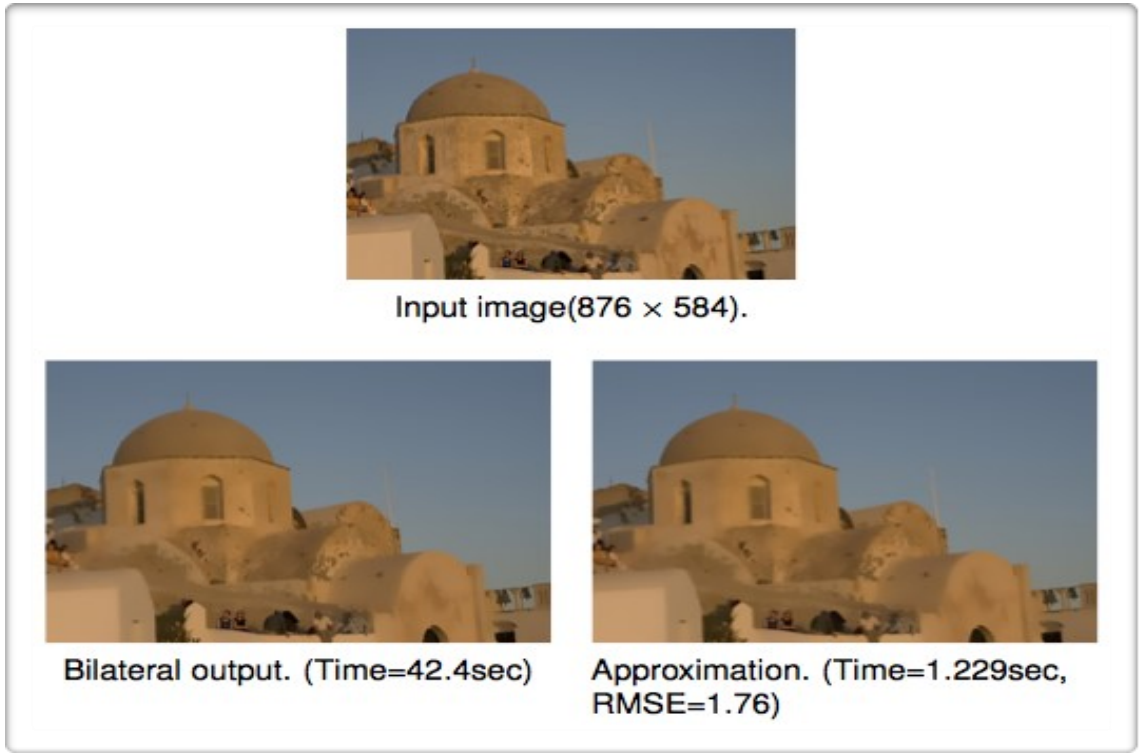


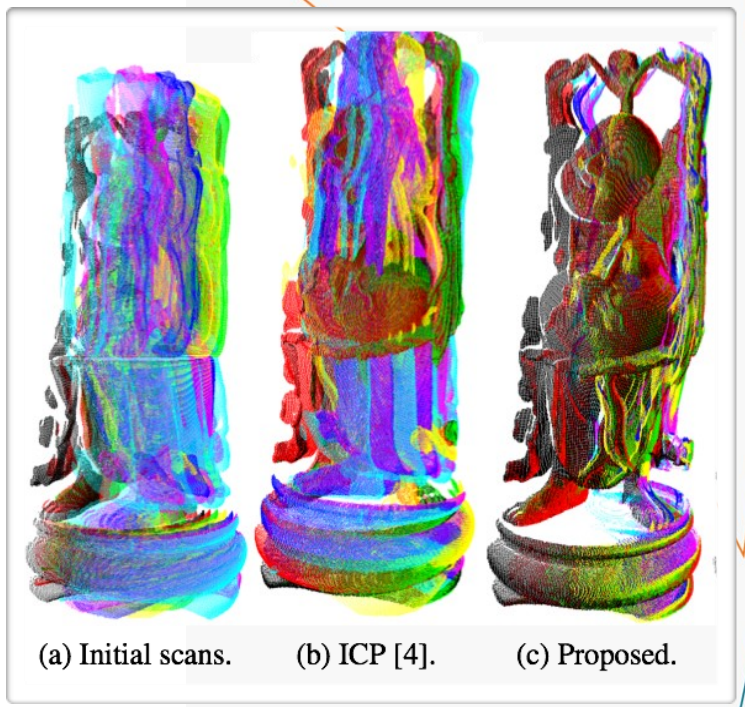
Intensity augmented second-order TV (our contribution)

Improved regularization in photo-acoustic imaging

Fast and Scalable Algorithms

- Image denoising, deblurring, inverse problems
- Sensor network localization from range data
- Multi-view registration from range scans
- Large-scale convex optimization algorithms
- Fast approximations of nonlinear filters with provable accuracy guarantees

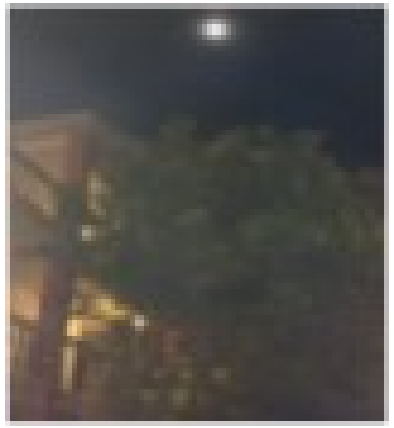


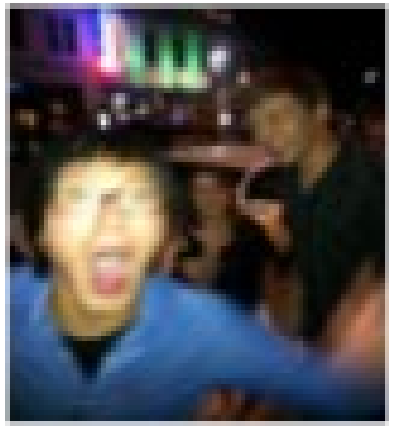


- Novel optimization algorithms and solvers for registration
- Applications in Computer Vision, Graphics, and Sensor network localization
- Novel combinatorial characterization of the rigidity (uniqueness) of registration over the Euclidean group

Image and Video Quality Assessment

- Virtual Reality – User experience
- Models of visual quality of 360 degree 3D images/videos
- Visual discomfort model (head mounted displays)
- Deliver better user experience
- Methods for camera captured images with authentic distortions
- Low light/improper exposure combined with noise and blur
- Perceptual trade-off in video streaming (rebuffering vs. compression)





amazon prime video

NETFLIX

hotstar

Document Image Processing & Assistive Technologies

Input

Mean character level accuracy is 25 %

Our DNN model designed to improve the quality of low-resolution degraded Tamil document images

Mean character level accuracy is 64 %

Output

- Online handwriting recognition system for Indian languages
- OCR systems for the blind
- Text-to-speech synthesis

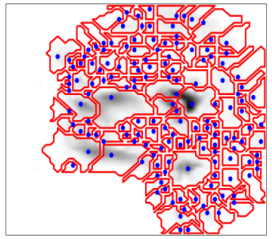
Speech Generation & Speech Processing

Deriving acoustic representations from real-time magnetic resonance imaging (rtMRI) videos of speech

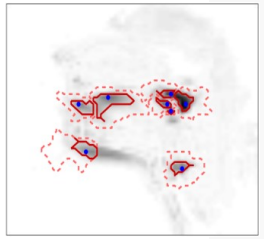
Information-theoretic vocal tract region selection

rtMRI Video

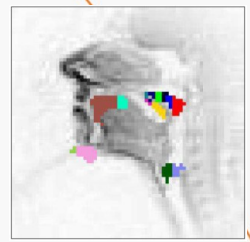
Pixel specific phonetic information (PSPI)



Pixel grouping



Optimal regions



Region splitting

$$\mathcal{I}(Z, X)$$

Mutual Information Pixel Value Sound class

- Local Maxima in PSPI
- $\mathcal{R}_d$ Region corresponding to d-th maxima

D #Regions of interest

η Threshold in a region

Pixel Intensity

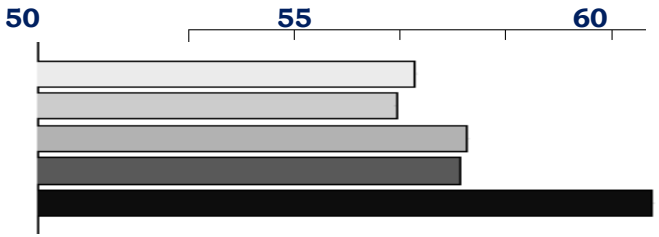
$$z_d = \frac{1}{|\bar{\mathcal{R}}_d|} \sum_{(i,j)\in\bar{\mathcal{R}}_d} \mathcal{M}(i,j)$$

$$\mathbf{z} = [z_1,\ z_2,\ \cdots,\ z_d,\ \cdots,\ z_D]^*$$

$$\{D^\star, \eta^\star\} = \arg\max_{D,\eta} \mathcal{I}(\mathbf{z}, X)$$

z_d - splitting

Phoneme Recognition Accuracy (%)



- Acoustic Features (ACF)
- Baseline Articulatory features (BAF)
- Optimal Articulatory features (OAF)
- ACF + BAF
- ACF + OAF

rtMRI derived articulatory features and acoustic features are complementary