August 1, 2016

Predicating What and Where in Social Media Images

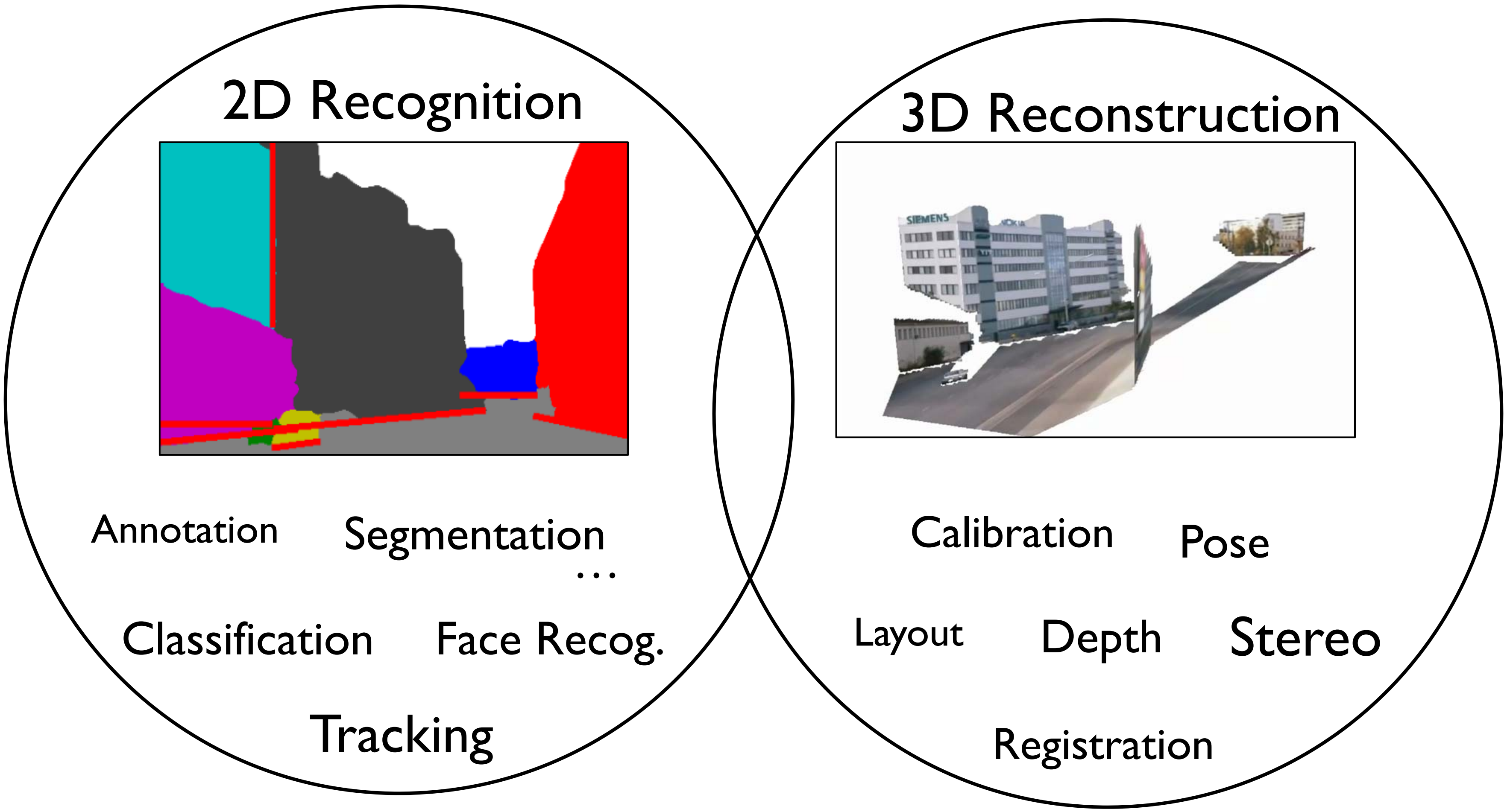
SAN DIEGO STATE UNIVERSITY

Xiaobai Liu,
Assistant Professor
Computer Vision Laboratory
Department of Computer Science
San Diego State University (SDSU)

- Email: xiaobai.liu@mail.sdsu.edu

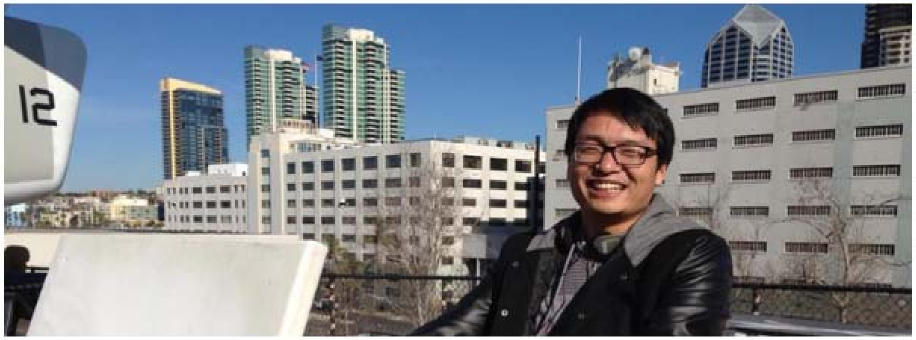
Computer Vision: What and Where

Computer Vision: What and Where



HTTP:// CV.SDSU.EDU

Liu Xiaobai , PhD



Assistant Professor,
Department of Computer Science,
San Diego State University, SD, 92182

Office: GMCS 542
Campus-Phone: 619-5944345
EMail: xiaobai Dot liu AT mail DOT sdsu DOT edu

News

July 22, 2016: Our paper on large-scale optimization will appear in the Journal TKDE (IF: 2.067).
July 10, 2016: Our paper on video coding will appear in the Journal FGCS (IF: 2.054).
June 25, 2016: I will chair the session: Maching Learning 8- Data Mining in IJCAI'2016. Welcome to Attend!
June 22, 2016: Our paper on Visual Vehicle-to-Vehicle (V3I) was accepted as Oral Presentation by ACM MM'2016
Apr. 10, 2016: Our paper on Hierarchical LSTM model for Scene Parsing was accepted by IJCAI'2016.
Apr. 10, 2016: Our paper on Mobile Landmark Search was accepted by IJCAI'2016.
Apr. 10, 2016: Our paper on Single-view 3D Scene Reconstruction was accepted by IJCAI'2016.
Feb. 29, 2016: Our Paper on Multi-view Human Tracking was accepted by IEEE CVPR'2016!
Feb. 24, 2016: Received the SDSU GREW Fellowship Spring 2016.
Feb. 1, 2016: I will chair the sessions of VIS: Pose Estimation and ML: Deep Learning I in AAAI' 2016. Welcome to Attend!
Jan. 10, 2016: Our Proposal to the SDSU Undergraduate Research Program has been awarded. Congradulations to Jacob Thalman!
Jan. 7, 2016: Received a donation of GPU K40 from the NVIDA Inc. Thanks NVIDA!
Dec. 1, 2015: Our paper on Attributed Grammar was accepted by AAAI' 2016.

Biography

I am working as Assistant Professor of Computer Science at the San Diego State University (SDSU) . I am also affiliated with the Center for Vision, Cognition, Learning and Autonomy (VCLA), University of Californiat, Los Angeles (UCLA).
In prior to joining SDSU, I worked as a Postdoctoral Research Scholar at the University of California, Los Angeles (UCLA) with Professor Song-Chun Zhu (from July 2013 to August 2015) and Professor Alan L. Yuille (from June 2011 to July 2013). I received my PhD degree from the Huazhong University of Science and Technology (HUST) in November, 2012. I was a visiting Doctoral Student at the National University of Singapore (NUS), Singapore, working Professor Shuicheng Yan from 2008-2011.

Teaching

Computer Vision & Machine Learning

Outline of this Talk

- **Part –I: 2D reconstruction (What)**
  (weakly supervised image parsing)
  - Label-to-Region
  - Label-to-region by search
  - Image Label Competition
  - Tree-structure sparsity
- **Part-II: 3D reconstruction (Where)**
  - Single-image Reconstruction (demo 1-3)
  - 3D Human tracking (demo 4-6)

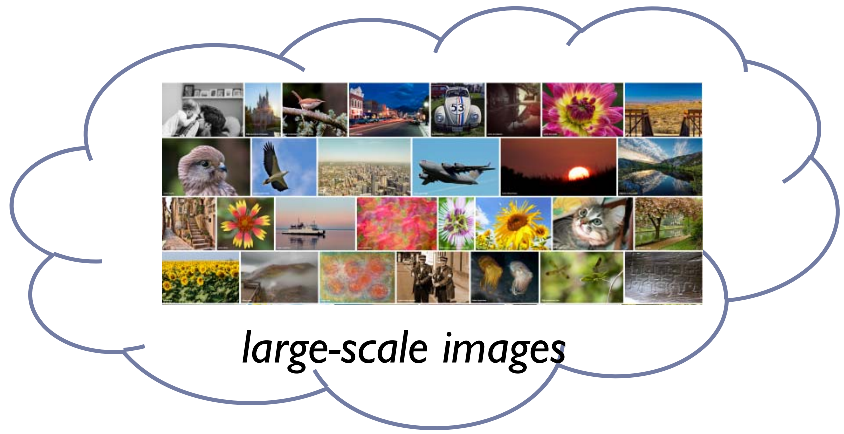
Nominated as one of two Best Paper Candidates in Content Track

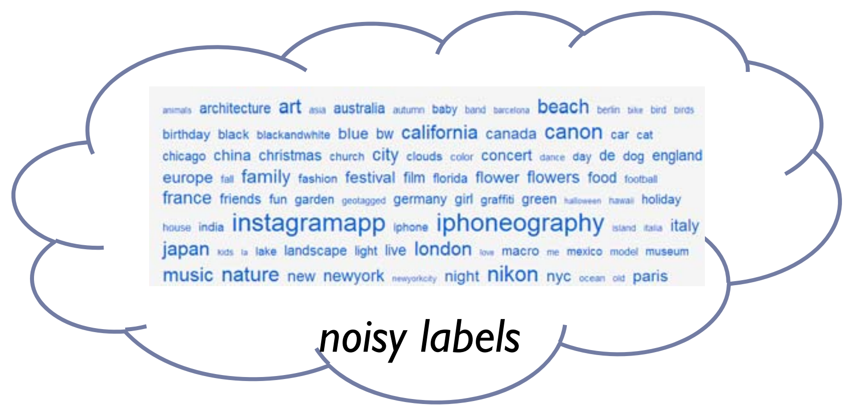
I. Label to Region by Bi-Layer Sparsity Priors

- **X. Liu**, B. Cheng, S. Yan, T. Chua, J. Tang and H. Jin., **Label to Region by Bi-Layer Sparsity Priors**. Proc. *ACM Conference on Multimedia* (***MM,*** Full Paper), 2009

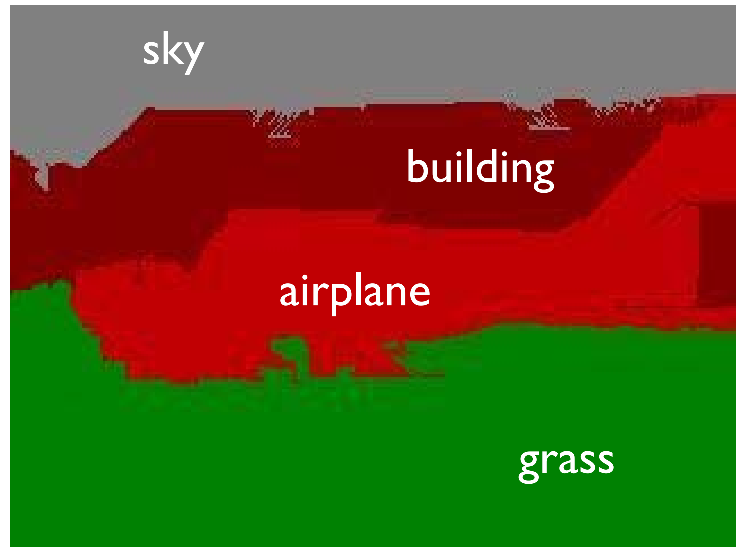
Online Photos

- Photo-sharing websites
  - ✓ Twitter
  - ✓ Flickr
  - ✓ Facebook
  - ✓ eBay
  - ✓ …
- Potentials
  - Content-based image retrieval
  - Visual Recommendation
  - Dr. Tsou's homepage



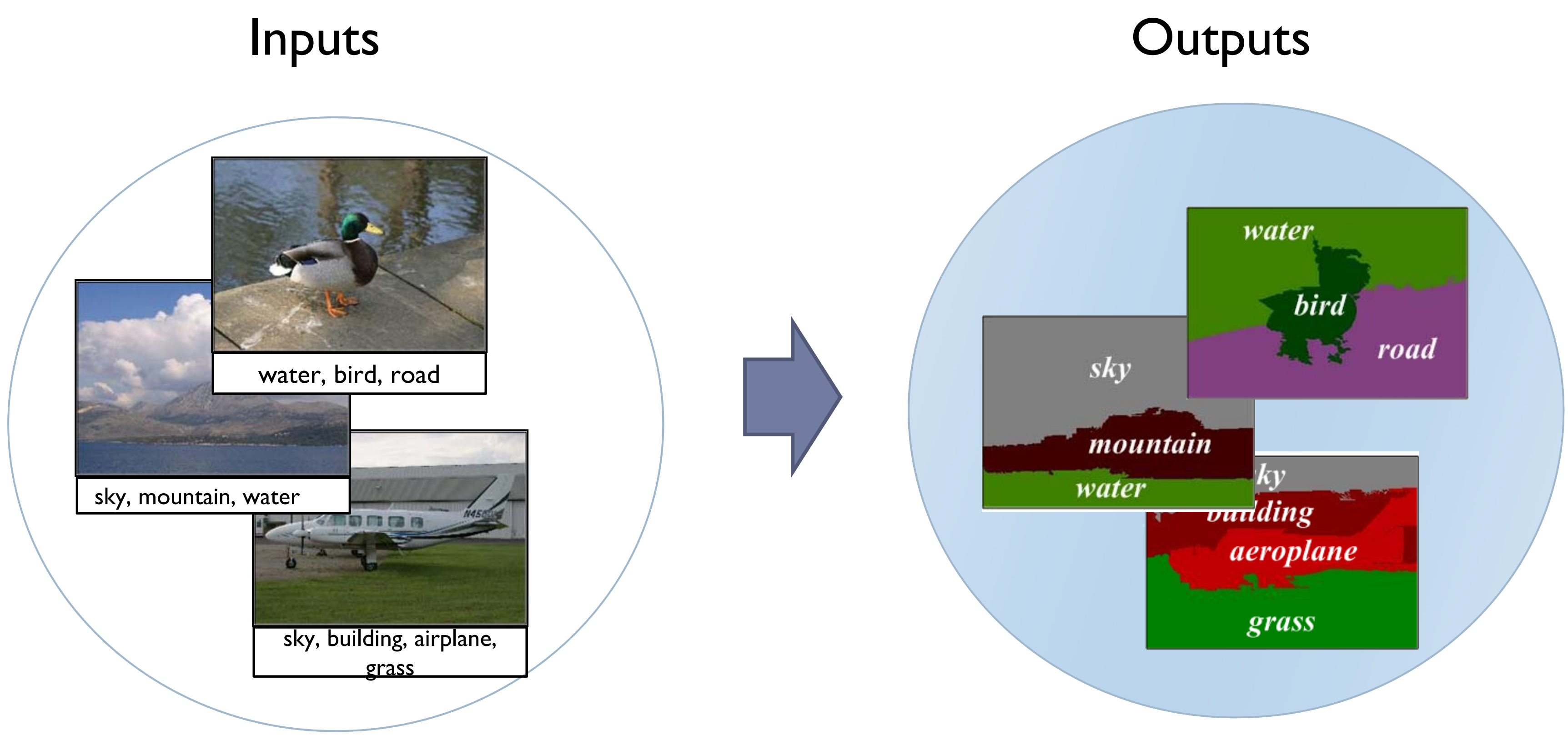


Task: Label to Region



Label to Region for a single Image is Challenging!

Task: Label to Region

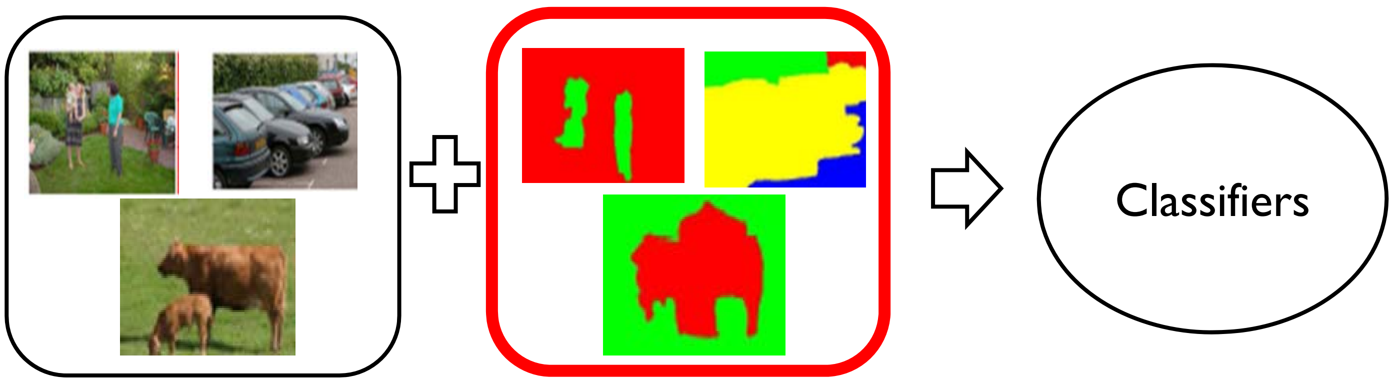


Simultaneous Region Partition and Labeling in Batch Mode

Related Work

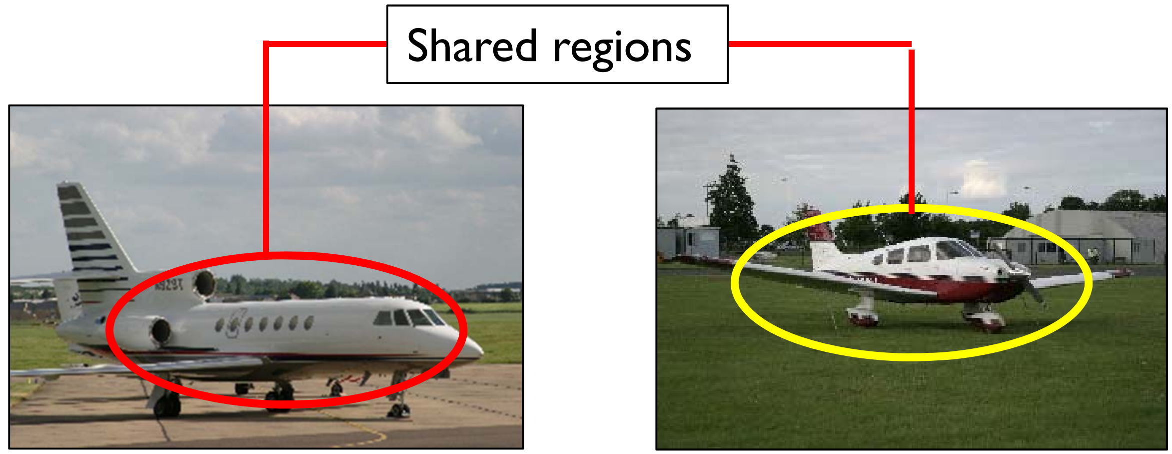
Supervised Learning Techniques

[C.Galleguillos et al., 2008][Jeon et al., 2003][Kang et al., 2006][Zhang et al., 2007]



Label-to-Region is valuable in Computer Vision community.

Label to Region: Our Approach



sky, road, aeroplane

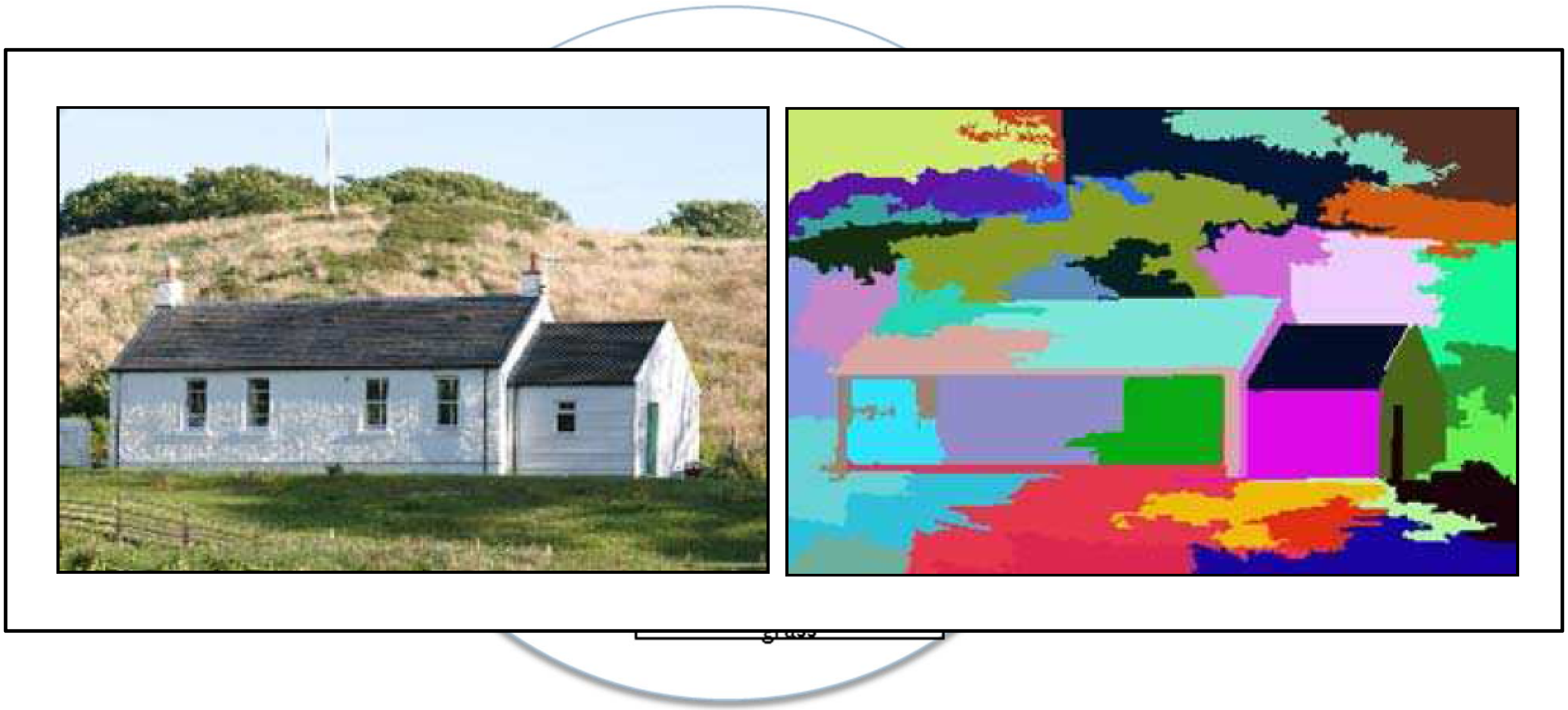
sky, grass, tree, aeroplane

Solution: for each pair of images, assign shared labels, if any, to shared regions!

Cross-Image Correspondence

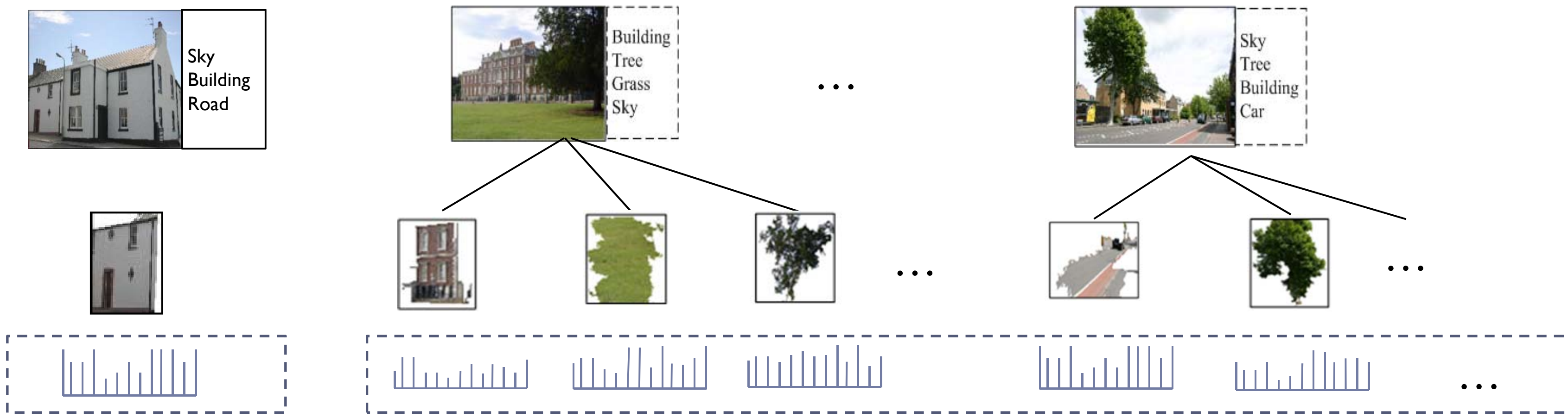
Label to Region: Correspondence

Step-1: Over-Segmentation



Label to Region: Our Approach

Step-2: cross-image correspondence



$$y = A_1\boxed{\alpha_1} + A_2\boxed{\alpha_2} + A_3\boxed{\alpha_3} + \cdots \quad A_4\boxed{\alpha_4} + \cdots$$

Using Coefficients as Relevance

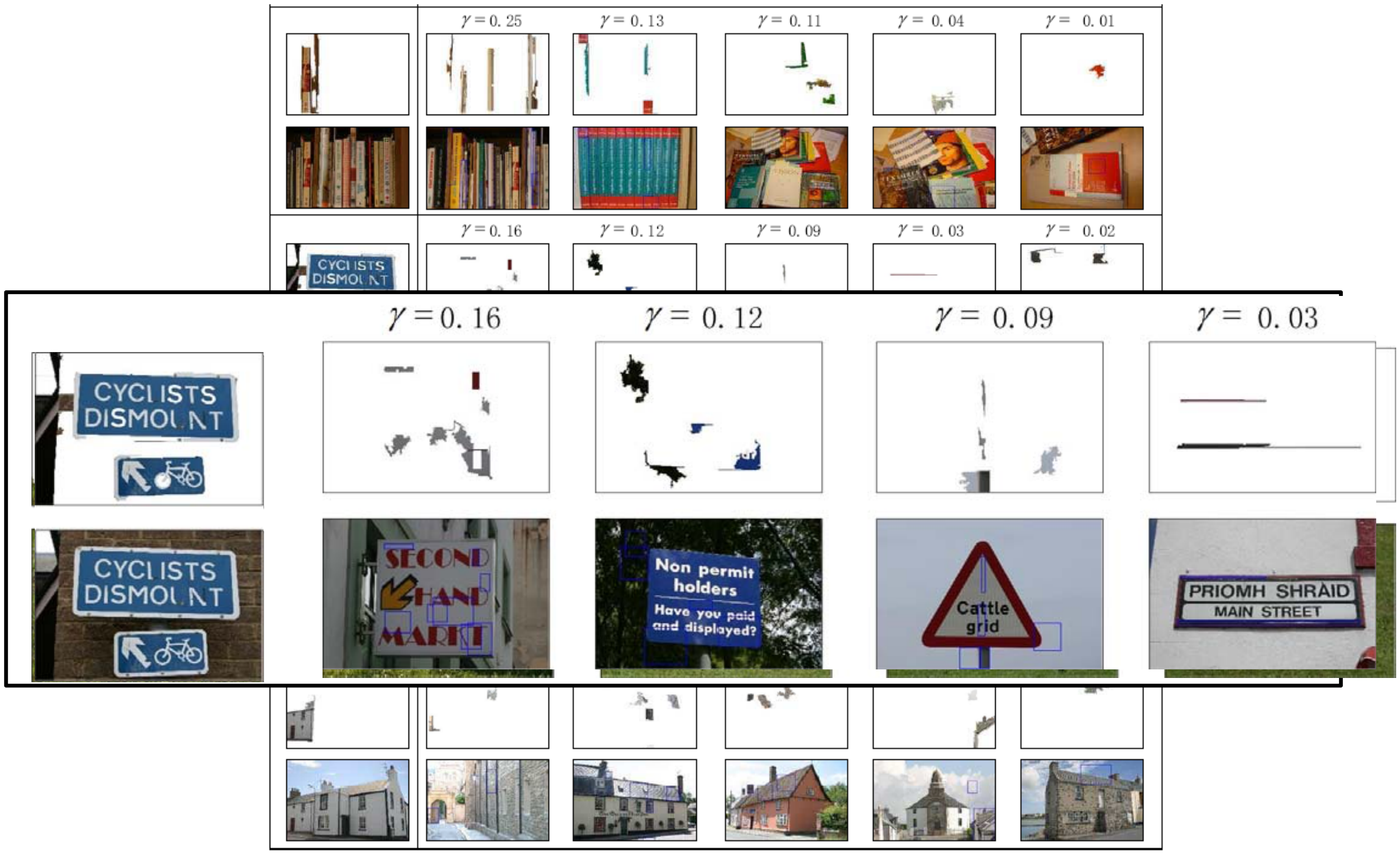
Criteria:

- *Select as few patches as possible;*
- *Select patches from as few images as possible:*

$$\arg\min_{\alpha,\epsilon,\gamma} ||\alpha||_1 + ||\epsilon||_1 + ||\gamma||_1 \quad s.t. \quad y = A\alpha + \epsilon,\ \gamma = B\alpha$$

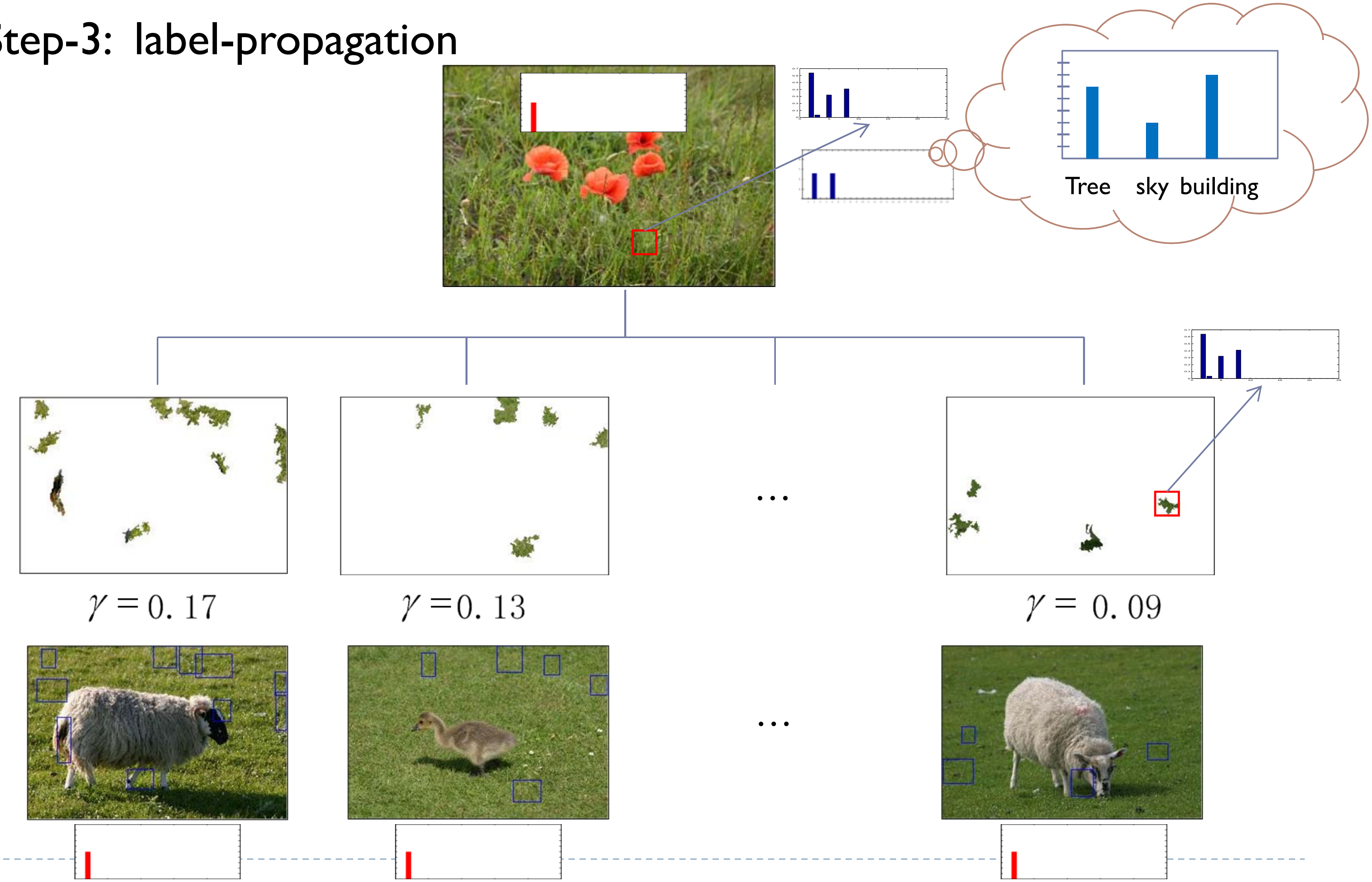
Bi-Layer Sparse Representation

Label to Region: correspondence

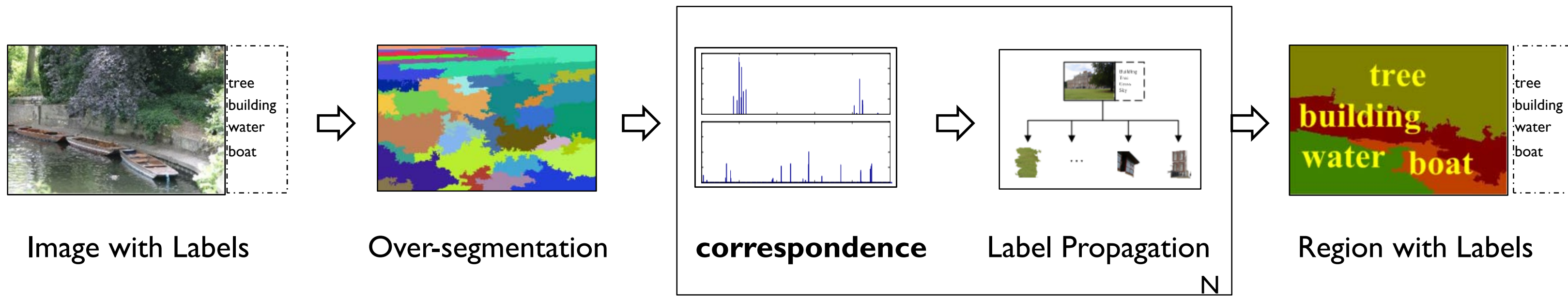


Label to Region： Label Assignment

Step-3: label-propagation



Label to Region: Our Approach



Step 1: ***Initialization***

- Over-segmentation

Step 2: Select one of the input images as reference image

Step 2.1: recover cross-image correspondence: bi-layer sparse representation

Step 2.2: bi-directional *label propagation*

Step 3: ***Post-Processing***

- Derive semantic regions and associated labels.

Label to Region: Results

MSRC dataset

Label to Region：Accuracies

Dataset	SVM-1	SVM-2	SVM-3	SVM-4	One-Layer	Bi-Layer
MSRC	0.22	0.20	0.24	0.23	0.47	0.63 0.81
COREL	0.29	0.32	0.33	0.32	0.51	0.61 0.76

The SVM-based algorithm is implemented with different values for the parameter of maximal patch size, namely, SVM-1: 150 pixels, SVM-2: 200 pixels, SVM-3: 400 pixels, and SVM-4: 600 pixels.

Summary of Label-to-Region

Contributions

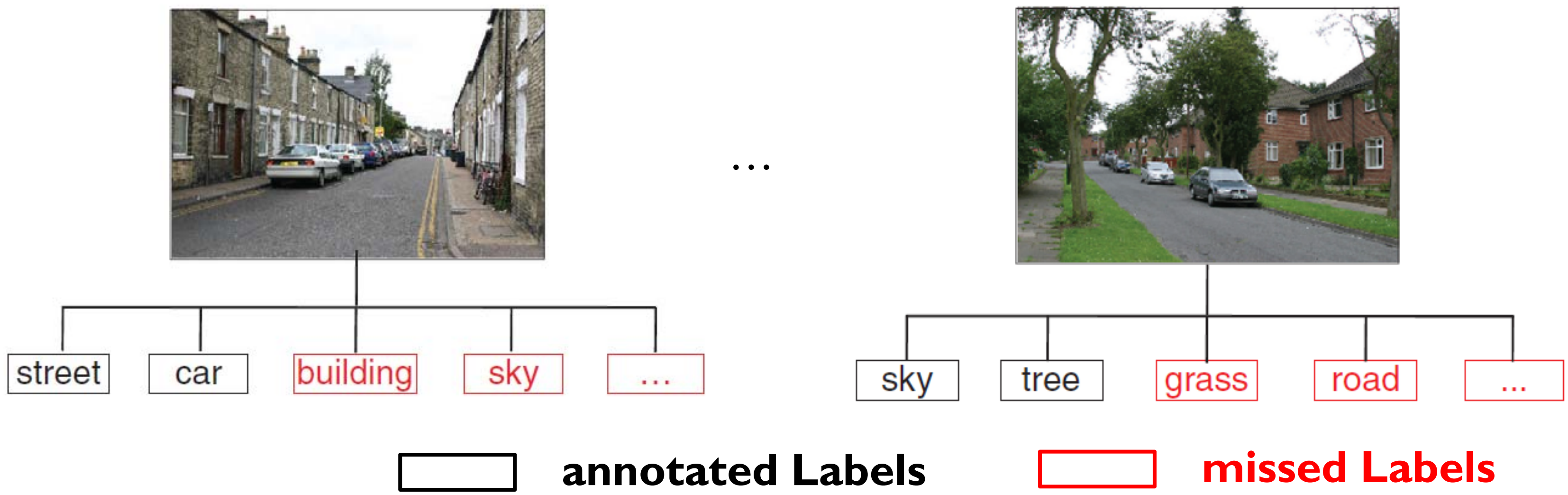
- Label-to-Region task
- Label propagation
- Bi-Layer sparsity Model

Limitations

- Can only handle labels corresponds with local region, e.g. , road;
- Process a set of images at the same time;
- Cannot handle partially annotated images or noisy tags;

II. Image Label Completion

Partially annotations or noisy labels

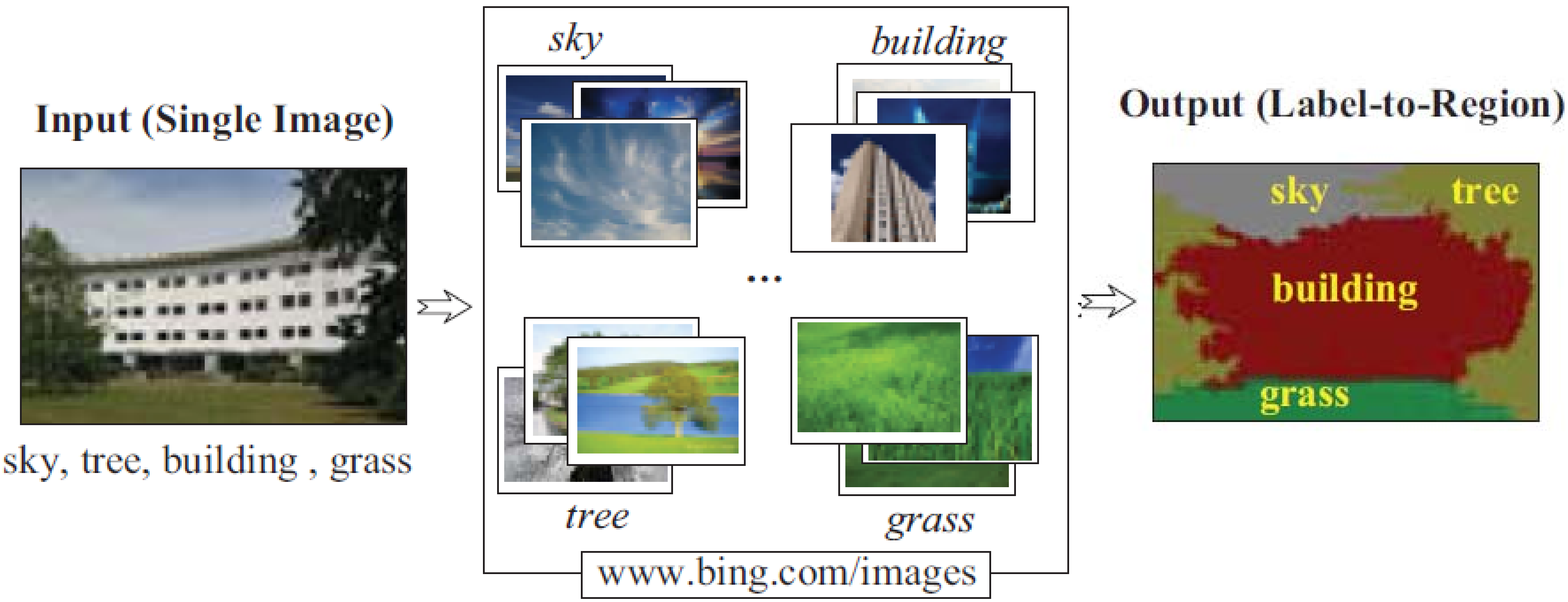


Label Completion via Nonnegative Decomposition

$$\min_{W,Y} \alpha Tr(WBW^T) + \beta Tr(CYLY^TC^T) + \gamma ||\tilde{Z}_0 \circ (CY)||^2 + ||X - WY||^2, \quad s.t.\ W, Y \geq 0,$$

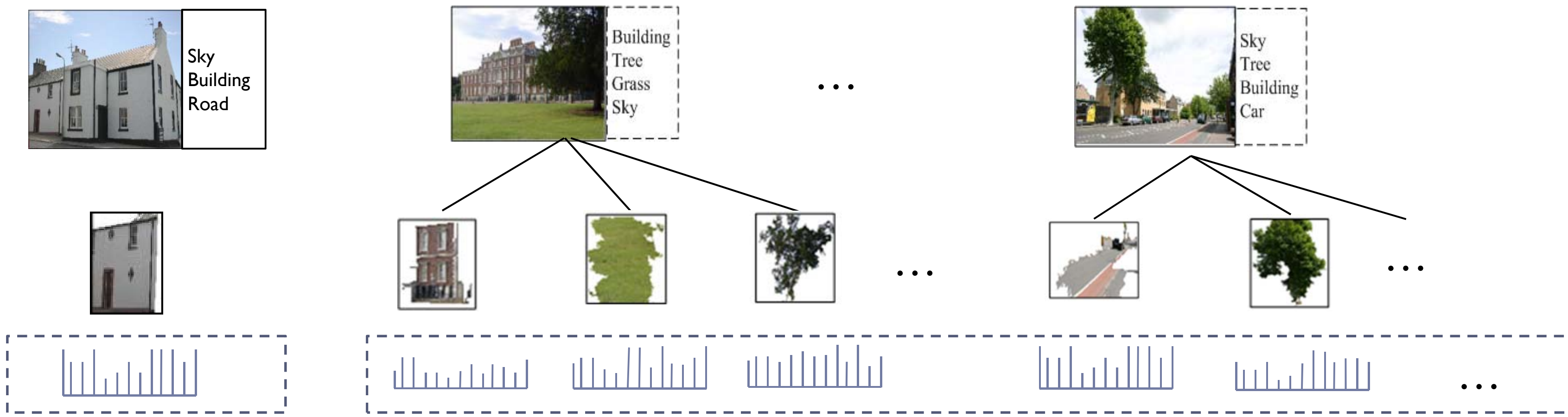
X. Liu, et al. IEEE Transactions on Image Processing, 2010

III. Label-to-region by Search



[Liu et al. IEEE CVPR'2010]

IV. Tree Structure Sparsity



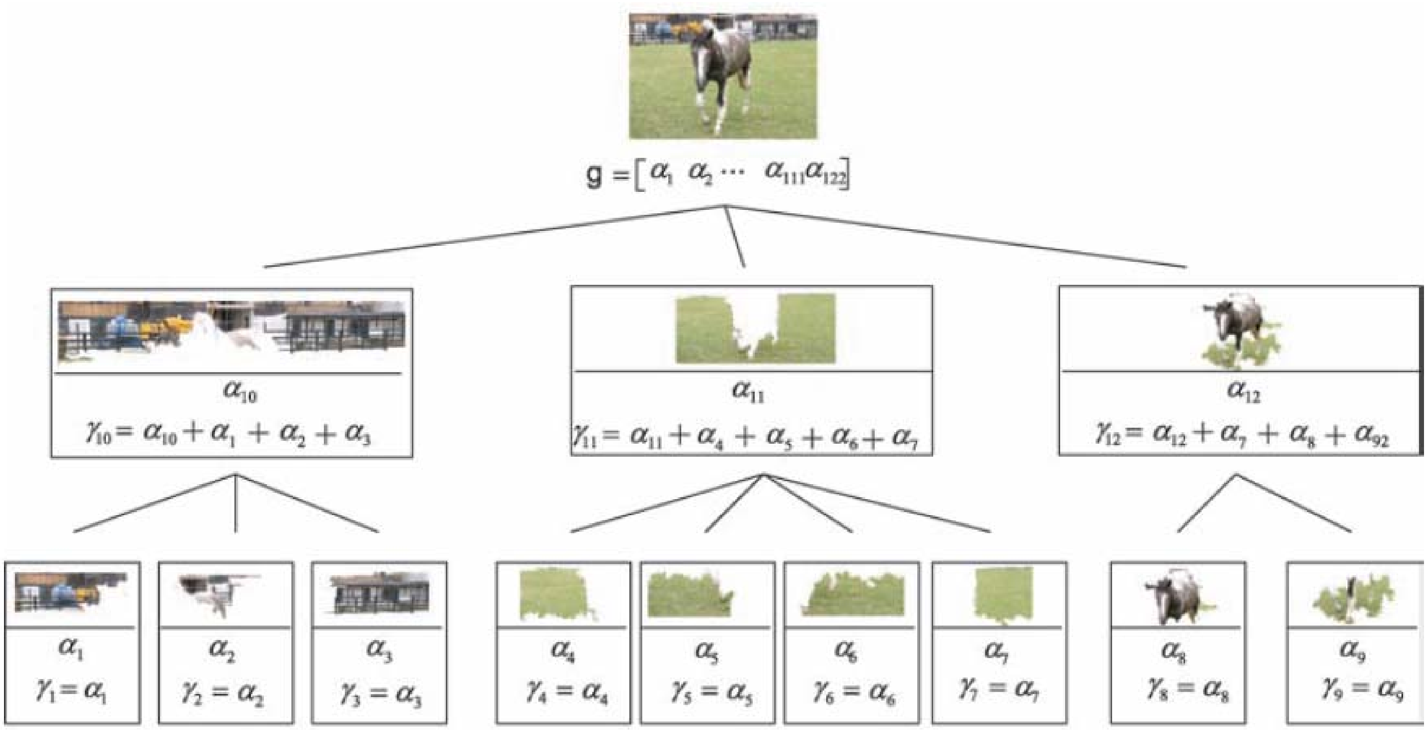
$$y = A_1\alpha_1 + A_2\alpha_2 + A_3\alpha_3 + \cdots \quad A_4\alpha_4 + \ldots$$

$$\arg\min_{\alpha,\epsilon,\gamma} ||\alpha||_1 + ||\epsilon||_1 + ||\gamma||_1 \quad s.t. \quad y = A\alpha + \epsilon,\ \gamma = B\alpha$$

Bi-Layer Sparse representation

[**X. Liu**, ACM Transaction on MCCAP 2012]

IV. Tree Structure Sparsity



From Bi-Layer to Tree Structure

[X. Liu, ACM Transaction MCCAP 2012]

Outline of this Talk

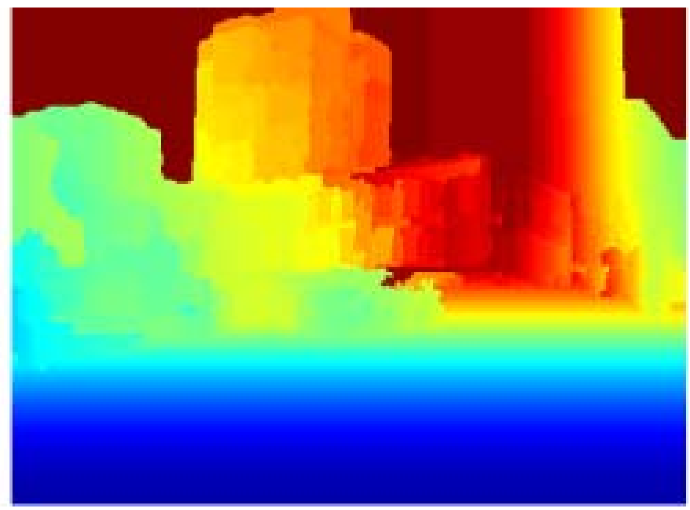
- **Part –I: 2D reconstruction (What)**
  (weakly supervised image parsing)
  - Label-to-Region
  - Label-to-region by search
  - Image Label Competition
  - Tree-structure sparsity
- **Part-II: 3D reconstruction (Where)**
  - Single-image Reconstruction (demo 1-3)
  - 3D Human tracking (demo 4-6)

Demo 1: 3D Scene Reconstruction from a single image

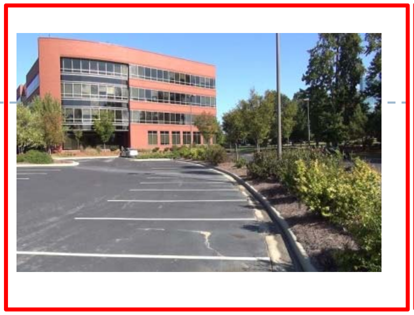
[Liu et al. CVPR'2014], [Liu et al. AAAI'2016], [Liu IJCAI'2016]

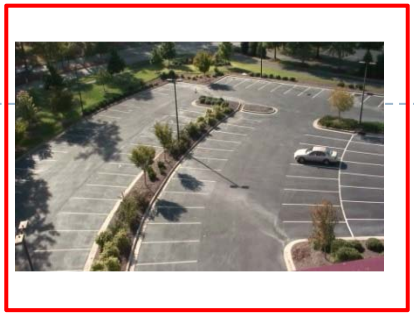
Demo 1: 3D Scene Reconstruction from a single image

Demo 1: 3D Scene Reconstruction from a single image

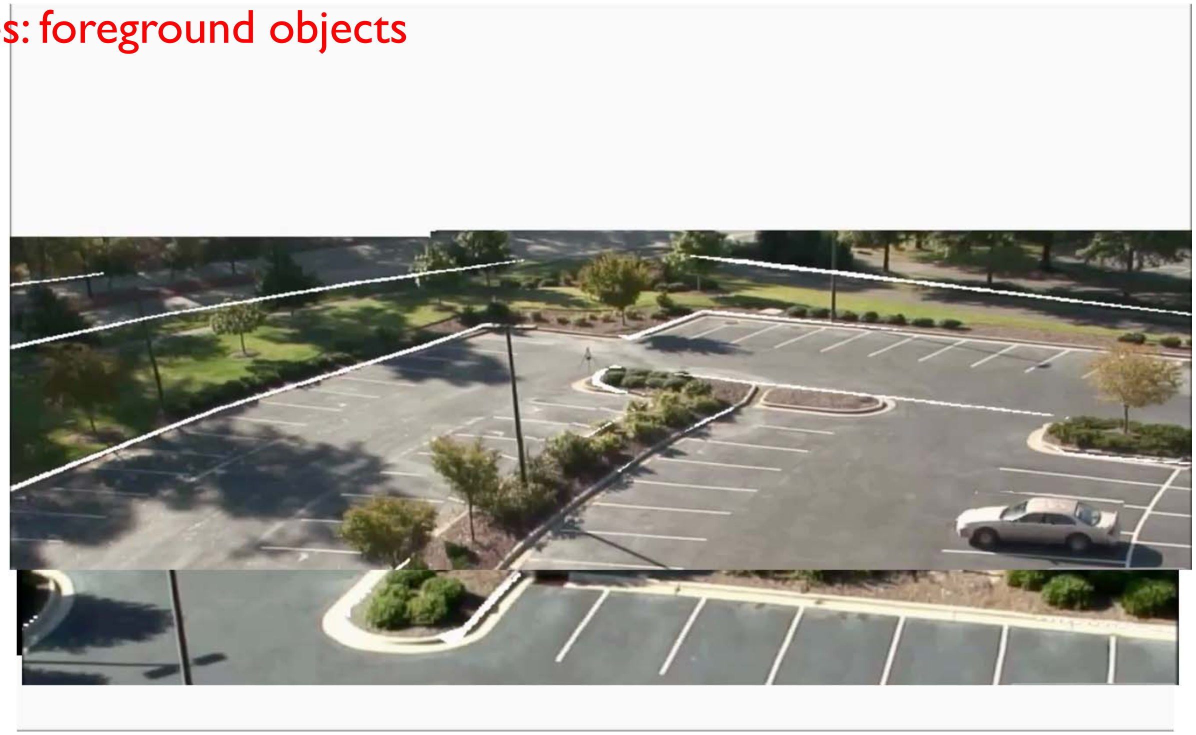


Demo 1: 3D Scene Reconstruction from a single image





Failure Cases: foreground objects



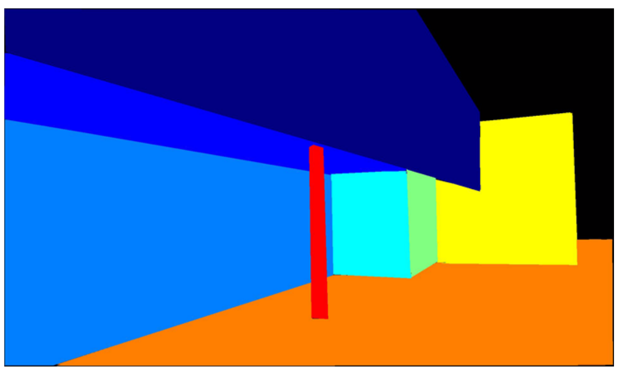
Synthesized Views

Demo 2: Scene Construction from Monocular Video

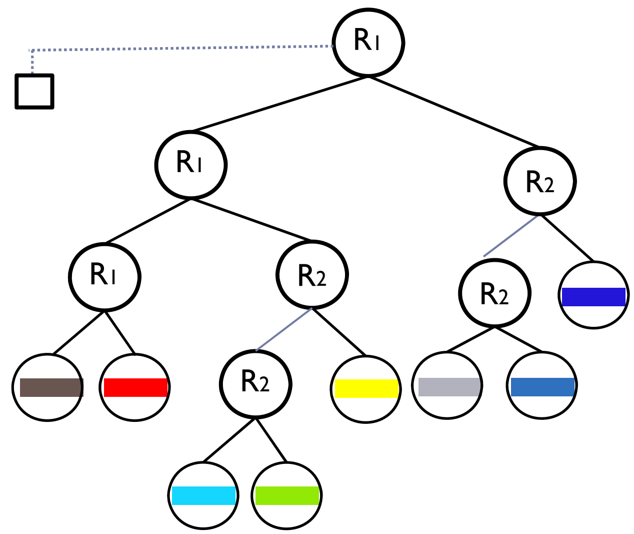
(captured by a mobile phone camera)

Input video sequence

Families of Parallel lines

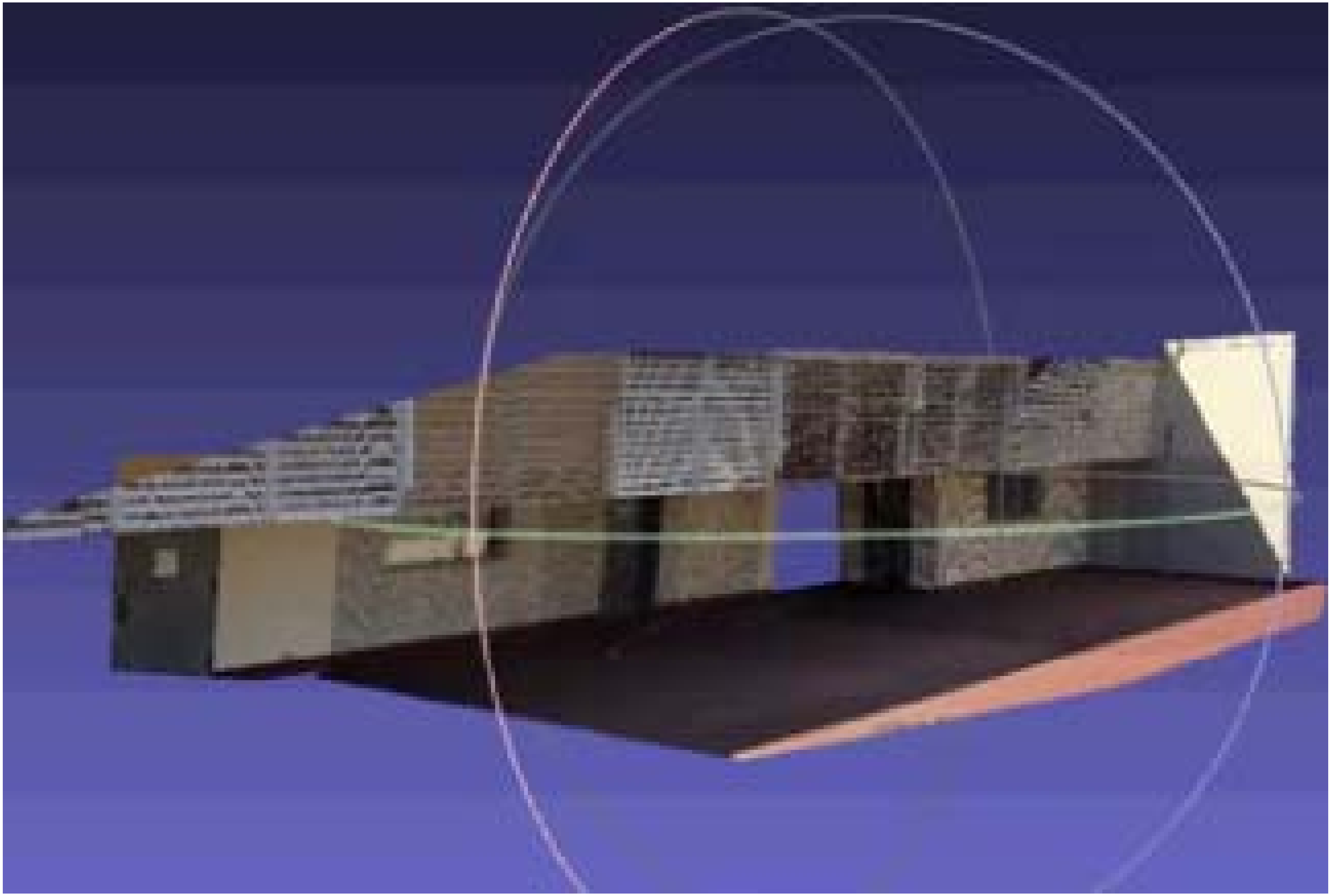


Layout Recovered

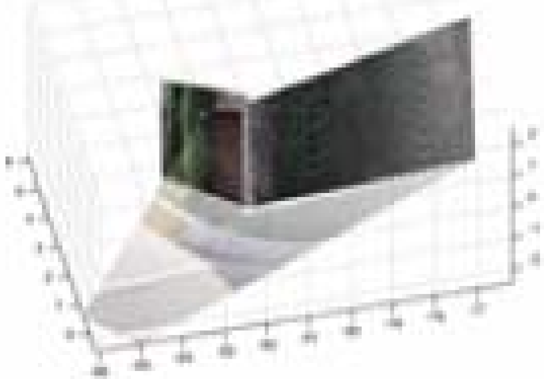


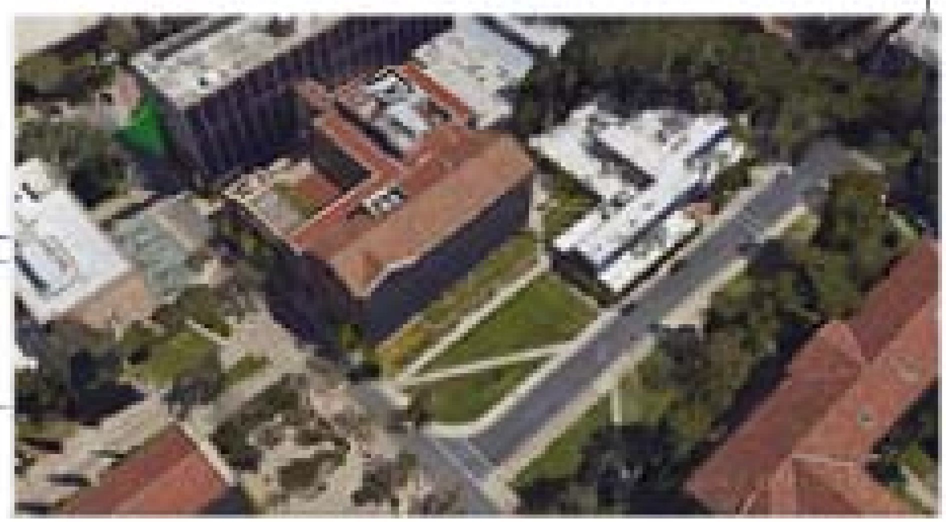
Notations	R_1: layout	R_3: affinity
R_2: siding	R_4: mesh	R_5: terminal

Recovered 3D



Demo 3: Scene Construction from Monocular Video

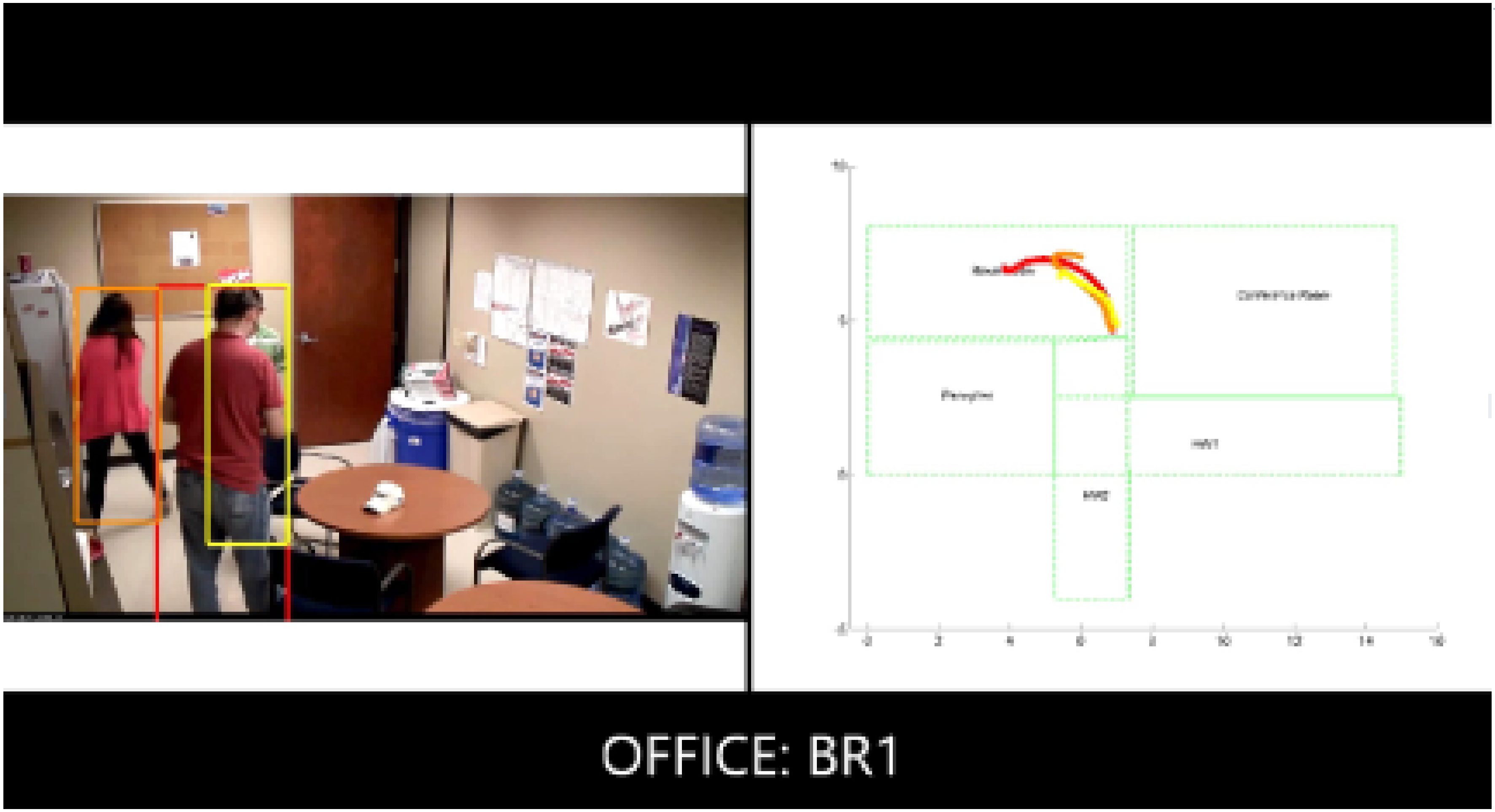




Outline of this Talk

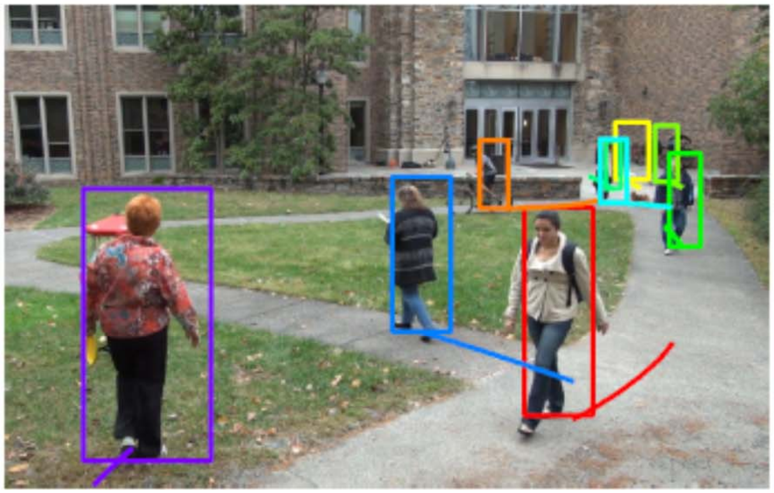
- **Part –I: 2D reconstruction (What)**
  (weakly supervised image parsing)
  - Label-to-Region
  - Label-to-region by search
  - Image Label Competition
  - Tree-structure sparsity
- **Part-II: 3D reconstruction (Where)**
  - Single-image Reconstruction (demo 1-3)
  - 3D Human tracking (demo 4-6)

Demo 4: Geo-Localization of moving persons (Indoor)



Liu et al. AAAI'2016

Demo 5: Geo-Localization of moving persons (outdoor)

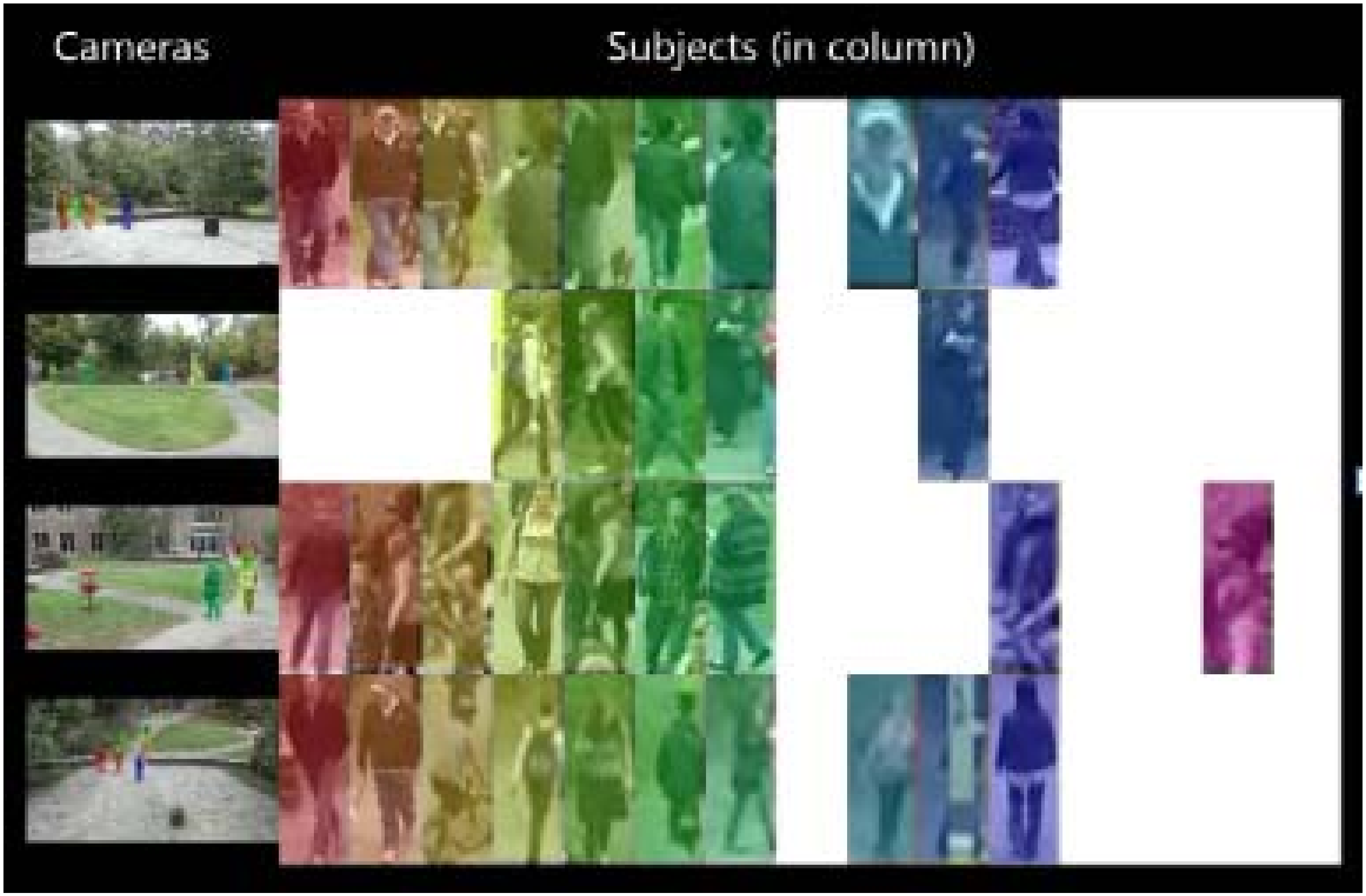


2D Trajectories

3D Trajectories

Liu et al. AAAI'2016

Demo 6: Multiple-View Multiple Person Tracking



Summary of this talk

- **Part –I: 2D reconstruction (What)**
  (weakly supervised image parsing)
  - Label-to-Region
  - Label-to-region by search
  - Image Label Competition
  - Tree-structure sparsity
- **Part-II: 3D reconstruction (Where)**
  - Single-image Reconstruction (demo 1-3)
  - 3D Human tracking (demo 4-6)

Question & Answer