Weakly Supervised Parsing of Web Images

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
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 lu AT mail DOT sdsu DOT edu

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 CREW Fellowship Spring 2016.
Feb. 1, 2016: I will chair the sessions of VIS: Pose Estimation and ML: Deep Learning I in AAAT 2016. Welcome to Attend!
Jan. 10, 2016: Our Proposal to the SDSU Undergraduate Research Program has been awarded. Congratulations to Jacob Thalman!
Jan. 7, 2016: Received a donation of GPU K40 from the NVIDIA Inc. Thanks NVIDIA!
Dec. 1, 2015: Our paper on Attributed Grammar was accepted by AAAT 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 California, 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

CS696: Applied Computer Vision, Spring, 2016 Syllabus
CS596: Machine Learning, Fall, 2015

Team Members

Grayson Atkins (Master’14, CS, SDSU)
Jingjie Yang (Master’14, CS, SDSU)
Shruthi Sriwah (Master’14, CS, SDSU)
Nithin Chakravarthy (Master’14, CS, SDSU)
Outline of this Talk

- Weakly supervised image parsing
  - Label-to-Region
  - Label-to-region by search
  - Image Label Competition
  - Tree-structure sparsity
I. Label to Region by Bi-Layer Sparsity Priors

Online Photos

- Photo-sharing websites
  - Flickr
  - Facebook
  - Twitter
  - eBay
  - …

- Potentials
  - Content-based image retrieval
  - Visual Recommendation
  - …

large-scale images

noisy labels
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

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

Using Coefficients as Relevance

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

\[
\text{arg min}_{\alpha, \epsilon, \gamma} ||\alpha||_1 + ||\epsilon||_1 + ||\gamma||_1 \quad \text{s.t.} \quad y = A\alpha + \epsilon, \gamma = B\alpha
\]

Bi-Layer Sparse Representation
Label to Region: Label Assignment

**Step-3: label-propagation**

- **$\gamma = 0.17$**
- **$\gamma = 0.13$**
- **$\gamma = 0.09$**
Label to Region: Results

MSRC dataset
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

\[
\min_{W,Y} \alpha \text{Tr}(W B W^T) + \beta \text{Tr}(C Y L Y^T C^T) + \gamma \| \tilde{Z}_0 \circ (C Y) \|^2 + \| X - W Y \|^2,
\]

\[s.t. \ W, Y \geq 0,\]

III. Label-to-region by Search

Input (Single Image)

sky, tree, building, grass

Output (Label-to-Region)

[Liu et al. IEEE CVPR’2010]
IV. Tree Structure Sparsity

\[
\gamma = A_1 \alpha_1 + A_2 \alpha_2 + A_3 \alpha_3 + \ldots + A_4 \alpha_4 + \ldots
\]

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

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]
Summary

- Weakly supervised image parsing
  - Label-to-Region (ACM ‘2009)
  - Label-to-region by search (IEEE CVPR’2010)
  - Image Label Competition (IEEE TIP’2010)
  - Tree-structure sparsity (ACM TOMCAP’2012)
Question & Answer