UNSUPERVISED DETECTION AND SEGMENTATION OF IDENTICAL OBJECTS

Minsu Cho*, Young Min Shin, Kyoung Mu Lee

Computer Vision Lab.
Dept. of EECS, Seoul National University, Korea
Homepage: http://cv.snu.ac.kr
Motivation

- How to discover **matching objects**?
  - FROM a single or multiple images
  - WITHOUT any supervision or specific object model
**Common Critical Limitation**

**ONE-TO-ONE object matching ACROSS two images**

- Not satisfied in real environments
- Still require some degree of supervision
Common Critical Limitation

ONE-TO-ONE object matching ACROSS two images
- Not satisfied in real environments
- Still require some degree of supervision
Motivation what if?

- Object matching in a SINGLE image?
Motivation what if?

- Many-to-Many object matching across an image PAIR?
Motivation what if?

- Many-to-Many-to-Many object matching across an image TRIPLET? …
Our Contribution

- **UNSUPERVISED** detection, segmentation, and grouping of identical objects from a single or multiple images

- Direct object discovery from images

  ‘Object Correspondence Networks’

  - Each network represents a set of identical objects
Our Contribution

- UNSUPERVISED detection, segmentation, and grouping of identical objects from images, even from a single

- Unsupervised image matching BEYOND ONE-TO-ONE object matching constraints

- Advanced image exploration in a more general and efficient match-growing framework than the previous methods

Cho et al., ECCV2008
Kannala et al, CVPR2008
Ferrari et al., IJCV 2006
A Brief Overview

- Problem decomposition
- Initial local seed matching
- **Multi-Layering**
- On a Bayesian formulation, augment the layers by
  - Intra-Layer Expansion
  - Inter-Layer Merge
- Networking reliable layers
Two types of object matching given images
- within a single image or across two images

Given $N$ images, $N$ intra- + $\frac{N(N-1)}{2}$ inter- matching problems
Initial Matching

- Local affine feature detector and descriptors
- Putative matches for our problem?
  - Conventional Nearest-Neighbor (NN) method
    - Lose inliers in the presence of multiple similar patterns
  - Allow multiple matches for each local feature
    - Using a loose similarity threshold in our method
    - Preserve true matches that do not belong to NNs
- Problem: drastic increase in outliers
Initial Matching

- MULTIPLE object matches against these OUTLIERS?

Solution: Exploring the images by

Multi-layer match-growing + Bayesian model
Each initial match builds up a singleton cluster with its own expansion layer.

- The layer provides the match cluster with space for expansion.
Multi-Layer Match-Growing

- The match clusters are iteratively grown by
  - Intra-layer Expansion: Propagating reliable matches into neighbors on the layer
  - Inter-layer Merge: Combining compatible matches of different layers

- Guided by our Bayesian model of object correspondences
Our target $\theta$ consists of object correspondences

$$\theta = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_K\}$$

$\Gamma_i$ : $i$th match cluster (object correspondence)

$$\Gamma_i = \{M_{i;m} : m = 1, \ldots, N_i\}$$

$M_{i;m}$ : $m$th match of $i$th cluster

$$M_{i;m} = \{(R_{i;m}^H, R_{i;m}^T), H_{i;m}\}$$

$H_{i;m}$ : homography between $R_{i;m}^H$ and $R_{i;m}^T$
The goal is to find $\theta^*$ that maximizes its posterior

$$\theta^* = \arg \max_{\theta} p(\theta \mid I) = \arg \max_{\theta} p(I \mid \theta) p(\theta)$$

- Photometry
- Geometry
- Maximality

Similarity of corresponding local appearance

Similarity of neighboring structure

Regularization for growing
Multi-Layer Match-Growing

- Multi-Layering
  - $U \sim [0,1] < Q_e$
    - Yes: Expansion Proposal
    - No: Merge Proposal
  - $p(\Theta | I)$
    - Converged:
      - Increased:
        - Update cluster and its expansion layers
        - Update the neighborhood system
      - Networking
Multi-Layer Match-Growing

- Expansion moves

- Expansion Proposal
  - Select a cluster
  - Select its member match as a supporter
  - Select a latent region around the supporter
  - Propagate it by the supporter, and refine
    Ferrari et al. IJCV06

- Initial expansion layer full overlapping regions
  Left: regions
  Right: their centers

- Reduction of its expansion layer along an expansion move

- Expansion proposal of a cluster
Multi-Layer Match-Growing

- Merge moves
  - Select two neighboring clusters
  - Combine two clusters
  - Merge proposal of two clusters
    - Their expansion layers are combined into intersection along the merge move

- Key design concept
  - Both expansion and merge moves gradually reduce the layers and guide each other to concentrate on potential region cooperatively.
Networking

- After convergence, connect the reliable match clusters (=layers)
Networking

- **Reliable match clusters: larger & textured**
  - Eliminate small or textureless ones based on the expanded size and intensity variance

- **Connecting match clusters**
  - Affinity of match clusters as their region overlap ratio
  - Agglomerative Clustering (single-link, cutoff: 0.8)

![Elimination of unreliable layers](image1)
![Connecting reliable layers](image2)
Review Demo

Input Image Pair
Review Demo

Decomposition and Initial Matching

Intra matching

Inter matching
Review Demo

Multi-Layer Match-Growing
Review Demo

Networking Reliable Object Correspondences
Experiments

Experimental settings

- For Initial matching
  - MSER & Harris-affine detector, SIFT descriptor
  - Collect the matches based on SIFT distance
  - Distance threshold 0.4 (max 1000 matches)

- Iterative Match Growing
  - $Q_e = 0.9$, $\Lambda_p = 2.0$

- Criteria for final valid clusters
  - Region expansion $> 2\%$ of the image area
  - Mean intensity variance $> 0.005$
Experiments #1
Unsupervised Object Discovery

- Dataset of images where several kinds of identical objects appear with clutter and occlusion
Experiments #1
Unsupervised Object Discovery

- Result examples
Experiments #1
Unsupervised Object Discovery
Experiments #1
Unsupervised Object Discovery
Experiments #1
Unsupervised Object Discovery

- Pair-wise Object Matching Accuracy

![Image of object matching accuracy with true negatives and false positives marked]

Recall: \( \frac{2}{3} = 0.67 \)
Precision: \( \frac{2}{3} = 0.67 \)

<table>
<thead>
<tr>
<th></th>
<th>Single</th>
<th>Pair</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall</td>
<td>0.89</td>
<td>0.61</td>
<td>0.75</td>
</tr>
<tr>
<td>Precision</td>
<td>0.94</td>
<td>0.89</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Average Recall and Precision of Pair-wise Object Matching on our dataset (10 singles, 10 pairs, total 30 images)
Experiments #2
On a Multi-Shot Image

- A penta-tiled image from ETHZ toys dataset

- Among 8 sets of identical objects
  - 1 object missing, 1 false positive
Experiments #3
Comparison with previous one-to-one methods

- Test on SNU dataset [http://cv.snu.ac.kr/~corecognition/]
  - Accuracy on common object detection & segmentation

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Mickey’s</th>
<th>Minnie’s</th>
<th>Jigsaws</th>
<th>Toys</th>
<th>Books</th>
<th>Bulletins</th>
<th>Our average</th>
<th>Average of [1]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hit Ratio</td>
<td>84.4%</td>
<td>87.8%</td>
<td>81.6%</td>
<td>85.2%</td>
<td>97.2%</td>
<td>93.8%</td>
<td><strong>88.3%</strong></td>
<td>85.5%</td>
</tr>
<tr>
<td>Bk Ratio</td>
<td>19.4%</td>
<td>31.1%</td>
<td>19.4%</td>
<td>20.8%</td>
<td>11.3%</td>
<td>15.2%</td>
<td><strong>19.5%</strong></td>
<td>22.4%</td>
</tr>
</tbody>
</table>

Cover the deformed part more accurately and completely

Our method is more accurate as well as more general
Experiments #4

Applications

- Automatic object-based 3D reconstruction
  - Multi-views of recurrent objects from a single view

A single input image

Identical object detection and segmentation

Object 3D reconstruction by multi-view stereo

We used Furukawa’s software for multi-view stereo

http://grail.cs.washington.edu/software/pmvs/pmvs-1/
Conclusion

- Fully unsupervised detection, segmentation, and grouping of identical objects based on an advanced match-growing approach
- More general in matching constraints
- More improved in accuracy

- A variety of potential applications
  - Object-based image retrieval
  - Automatic object modeling & reconstruction
  - Image analysis and compression
  - Etc.
Thank you

http://cv.snu.ac.kr/~minsucho