Salient object detection, objectness proposals, and potential applications

Slides: Ming-Ming Cheng
Presenter: Jian Li

Tutorial webpage: http://mmcheng.net/saltutorial/
Predict fixation $\Rightarrow$ detect salient object

- Fixation prediction
  - Predicting saliency points of human eye movement

- A model of saliency-based visual attention for rapid scene analysis. PAMI 1998, Itti et al.
- Saliency detection: A spectral residual approach. CVPR 2007, Hou et al.
- Graph-based visual saliency. NIPS, Harel et al.
Predict fixation $\rightarrow$ detect salient object

- Eye tracker
  - Cognitive psychology, neurobiology, etc.
Recent Advance in Salient object detection, objectness proposals, and potential applications

Predict fixation $\rightarrow$ detect salient object

• Saliency detection as binary segmentation

(a) MSRA10K  (b) ECSSD

- Learning to detect a salient object. CVPR 2007, Liu et. al.
- Frequency-tuned salient region detection, CVPR 2009, Achanta et. al.
- Global contrast based salient region detection, CVPR 2011, Cheng et. al.
Salient object detection

• How to define salient objects?
  • Match the human annotators’ behavior when they have been asked to pick a salient object in an image.

Global contrast based salient region detection, IEEE TPAMI 2015 (CVPR 2011), Cheng et al.
Salient object detection

• High consistency among labelers.

PASCAL-S dataset

The Secrets of Salient Object Segmentation, CVPR 2014, Li et. al.
Growing interest

What makes an object salient?

Hypothesis for Salient object detection

• Local contrast

Hypothesis for Salient object detection

- Global contrast

\[ S(r_k) = \sum_{r_k \neq r_i} \exp(-D_s(r_k, r_i)) w(r_i) D_r(r_k, r_i) \]

Regional contrast by sparse histogram comparison.

Global Contrast based Salient Region detection. IEEE TPAMI 2015 (CVPR 2011), Cheng et al.
Hypothesis for Salient object detection

• Spatial distribution

- Learning to Detect a Salient Object, IEEE TPAMI 2011 (CVPR 2007), Liu et al.
- Efficient Salient Region Detection with Soft Image Abstraction, ICCV 2013, Cheng et al.
Hypothesis for Salient object detection

- Sparse noises
  - Background residing in a low dimensional space with salient objects as sparse noises
Hypothesis for Salient object detection

• Focusness

Salient Region Detection by UFO: Uniqueness, Focusness and Objectness, ICCV 2013, Jiang et al.
Hypothesis for Salient object detection

- Center prior/bias

Hypothesis for Salient object detection

• Backgroundness

Geodesic saliency using background priors, ECCV, 2012, Wei et al.
Hypothesis for Salient object detection

• Objectness

Figure 1. An illustration of our approach from images to the final saliency map: (a) Input Image (b) objectness detections, (c) saliency prior from objectness, (d) diverse density scores for pixels, (e) the final saliency map, and (f) the segmented object.

Category-independent object-level saliency detection, ICCV 2013, Jia et. al.
Hypothesis for Salient object detection

• Convexity

Salient object detection using concavity context, ICCV 2011, Lu et al.
Recent advances in salient object detection, objectness proposals, and potential applications

ICIP 2015

1. Block-based vs. region-based
   - Pixels/patches $\to$ regions/super-pixels
   - Efficiency consideration
   - Abstract unnecessary details
   - Region contain complementary cues

### Table: Model Comparison

<table>
<thead>
<tr>
<th>#</th>
<th>Model</th>
<th>Pub</th>
<th>Year</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FG [10]</td>
<td>MM</td>
<td>2003</td>
<td>PI</td>
</tr>
<tr>
<td>2</td>
<td>RSA [77]</td>
<td>MM</td>
<td>2005</td>
<td>PA</td>
</tr>
<tr>
<td>3</td>
<td>RE [31]</td>
<td>ICME</td>
<td>2006</td>
<td>mPI + RE</td>
</tr>
<tr>
<td>4</td>
<td>RU [85]</td>
<td>TMM</td>
<td>2007</td>
<td>RE</td>
</tr>
<tr>
<td>5</td>
<td>AC [9]</td>
<td>ICVS</td>
<td>2008</td>
<td>mPA</td>
</tr>
<tr>
<td>6</td>
<td>FT [81]</td>
<td>CVPR</td>
<td>2009</td>
<td>PI</td>
</tr>
<tr>
<td>7</td>
<td>ICC [80]</td>
<td>CVPR</td>
<td>2009</td>
<td>PI</td>
</tr>
<tr>
<td>8</td>
<td>EDS [79]</td>
<td>PR</td>
<td>2009</td>
<td>PI</td>
</tr>
<tr>
<td>9</td>
<td>CSM [55]</td>
<td>MM</td>
<td>2010</td>
<td>PI + PA</td>
</tr>
<tr>
<td>10</td>
<td>RC [86]</td>
<td>CVPR</td>
<td>2011</td>
<td>RE</td>
</tr>
<tr>
<td>11</td>
<td>HC [96]</td>
<td>CVPR</td>
<td>2011</td>
<td>RE</td>
</tr>
<tr>
<td>12</td>
<td>CC [96]</td>
<td>ICCV</td>
<td>2011</td>
<td>mRE</td>
</tr>
<tr>
<td>13</td>
<td>CSD [82]</td>
<td>ICCV</td>
<td>2011</td>
<td>mPA</td>
</tr>
<tr>
<td>14</td>
<td>DSO [73]</td>
<td>ICCV</td>
<td>2011</td>
<td>mRE</td>
</tr>
<tr>
<td>15</td>
<td>LEO [97]</td>
<td>BMVC</td>
<td>2011</td>
<td>mRE</td>
</tr>
<tr>
<td>16</td>
<td>SF [54]</td>
<td>CVPR</td>
<td>2012</td>
<td>RE</td>
</tr>
<tr>
<td>17</td>
<td>SRE [90]</td>
<td>CVPR</td>
<td>2012</td>
<td>RE</td>
</tr>
<tr>
<td>18</td>
<td>TIP [67]</td>
<td>TIP</td>
<td>2013</td>
<td>hRE</td>
</tr>
<tr>
<td>19</td>
<td>MMR [101]</td>
<td>CVPR</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>20</td>
<td>IVC [94]</td>
<td>CVPR</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>21</td>
<td>FPA [84]</td>
<td>CVPR</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>22</td>
<td>GU [89]</td>
<td>ICCV</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>23</td>
<td>SRE [89]</td>
<td>ICCV</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>24</td>
<td>OA [100]</td>
<td>ICCV</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>25</td>
<td>BDR [109]</td>
<td>CVPR</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>26</td>
<td>MC [103]</td>
<td>ICCV</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>27</td>
<td>UFO [104]</td>
<td>ICCV</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>28</td>
<td>CIO [105]</td>
<td>ICCV</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>29</td>
<td>SLMR [106]</td>
<td>BMVC</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>30</td>
<td>LSM [107]</td>
<td>AAAI</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>31</td>
<td>SUB [90]</td>
<td>ICCV</td>
<td>2013</td>
<td>RE</td>
</tr>
<tr>
<td>32</td>
<td>PDE [108]</td>
<td>CVPR</td>
<td>2014</td>
<td>RE</td>
</tr>
<tr>
<td>33</td>
<td>RBD [109]</td>
<td>CVPR</td>
<td>2014</td>
<td>RE</td>
</tr>
</tbody>
</table>

Constantly shifting to region-based analysis for robustness.

Intrinsic cues vs. Extrinsic cues

• Intrinsic cues only from input image itself
  • Contrast, spatial distribution, center prior, etc.

• Extrinsic cues of similar images
  • User annotations, depth map, or statistical information of similar images
Heuristics vs. feature learning

• Use supervised learning approach to map the regional feature vector to a saliency score

Figure 3. The most important 20 regional features. See Table 1 and Table 2 for the description of the features.

Salient Object Detection: A Discriminative Regional Feature Integration Approach, CVPR 2013, Jiang et al.
Recent Advances in Salient object detection, objectness proposals, and potential applications

ICIP 2015

Saliency maps aggregation/optimization

- Linear
- Non-Linear
- Adaptive
- Hierarchical
- Bayesian
- Energy minimization
- Least square solver
- Gaussian MRF

How to optimally combine different saliency maps is still an open problem.
Enables automatic initialization provided by salient object detection.

- Global Contrast based Salient Region detection. IEEE TPAMI 2015 (CVPR 2011), Cheng et al.
Applications

• Is salient object detection for ‘simple’ images useful?

Applications

• Illustration of learned appearance models
• Accords with our understanding of these categories
Applications

Applications

Applications

Saliency for image Manipulation, The Visual Computer 2013, Margolin et al.
Applications

Applications

Unsupervised Object Discovery via Saliency-Guided Multiple Class Learning, IEEE CVPR 2012. Zhu et al.
Applications


Improve state-of-the-art by 10+% on PASCAL VOC!!
Applications: what we learnt?

Don’t ask what segments can do for you, ask what you can do for the segments.

— Jitendra Malik
How about complicated images?
Objectness proposals
Motivation: What is an object?
Motivation: What is an object?

• An objectness measure
  • A value to reflects how likely an image window covers an object of any category.

• What’s the benefits?
  • Improve computational efficiency, reduce the search space
  • Allowing the usage of strong classifiers during testing, improve accuracy

Measuring the objectness of image window, IEEE TPAMI 2012, Alexe et. al.
Motivation: What is an object?

• What is a good objectness measure?
  • Achieve **high object detection rate (DR)**
    • Any undetected objects at this stage cannot be recovered later
  • Produce a **small number of proposals**
    • Reducing computational time of subsequent detectors
• Obtain **high computational efficiency**
  • The method can be easily involved in various applications
  • Especially for realtime and large-scale applications;
• Have **good generalization ability** to unseen object categories
  • The proposals can be reused by many category specific detectors
  • Greatly reduce the computation for each of them.
A feature integration approach

• Objectness proposal generation
  • A small number (e.g. 1K) of category-independent proposals
  • Expected to cover all objects in an image

Region merging & Diversification

• Region merging
  • Merge two most similar regions based on region similarity.
  • Update similarities between the new region and its neighbors.

• Diversification

[Link: Selective Search for Object Recognition, IJCV 2013, Uijlings et. Al.]
Local & global search

• Local search
  • Unsuitable for object with distinct parts

• Global search
  • Initialize with foreground/background seeds
  • A global optimization function for each parameter set

Generating object segmentation proposals using global and local search, CVPR 2014, Rantalankila et al.
BING method

• Our observation: a small interactive demo
  • Take you pen and paper and draw an object which is current in your mind.
  • What if we resize it to a tiny fixed size?
    • E.g. 8x8. Not only changing the scale, but also aspect ratio.

BING: Binarized Normed Gradients for Objectness Estimation at 300fps, IEEE CVPR 2014 (Oral), M.M. Cheng, et. al.
BING method

- Objects are stand-alone things with well defined closed boundaries and centers.

- Using stuff to find things. ECCV 2008, Heitz et. al.
- Measuring the objectness of image window, IEEE TPAMI 2012, Alexe et. al.
Experimental results of BING method

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (seconds)</td>
<td>89.2</td>
<td>3.14</td>
<td>1.32</td>
<td>11.2</td>
<td><strong>0.003</strong></td>
</tr>
</tbody>
</table>

Table 1. Average computational time on VOC2007.
Combining boxes and regions

(a) Input  (b) Initial boxes  (c) Box Alignment  (d) $\delta = 0.7$  (e) $\delta = 0.3$

Improving Object Proposals with Multi-Thresholding Straddling Expansion, IEEE CVPR 2015, Chen, et. al.
Combining boxes and regions

- Experimental results

Improving Object Proposals with Multi-Thresholding Straddling Expansion, IEEE CVPR 2015, Chen, et. al.
Recent advance in salient object detection, objectness proposals, and potential applications

ICIP 2015

Applications

Future work in objectness proposals

• High detection rate under large IoU
• Running speed
• Small number of proposals
• Exploring more applications
Thanks!

Tutorial webpage: [http://mmcheng.net/saltutorial/](http://mmcheng.net/saltutorial/)