Large Scale Asset Extraction for Urban Images

Lama Affara, Liangliang Nan, Bernard Ghanem, Peter Wonka
King Abdullah University of Science and Technology (KAUST), Saudi Arabia

OVERALL PIPELINE

Preprocessing: Given an input image, we first detect line segments and rectify the image based on the detected most dominant facade. We then extract rectangular superpixels to restrict the search space for assets.

Prior Estimation: We employ a particle filter to estimate prior distribution functions (pdf) for the four bounding box parameters. For each parameter, three pdfs are estimated and updated during detection: global prior, history prior, and image prior.

Object Proposals: Object proposals are sampled as the evolving particle states and guided by the search space induced by estimated rectangular superpixels.

Asset Extraction: The proposed objects are classified, and the classifier’s output score is used to update the pdfs, and thus guide the sampled object proposals.

because assets might be captured from different viewpoints, their corresponding image regions should be rectified. We propose a novel image rectification algorithm that is 18 times faster than previous work. We breakdown the full rectification transformation into a concatenation of two simpler transformations: vertical perspective and horizontal perspective.

RECTIFICATION

GUIDED OBJECT PROPOSALS

We improve upon state-of-the-art object proposals by using the concept of interleaved proposing and classification.

The priors are updated using a particle filter strategy. The output of the classification at each time stamp is fed back as weights to the particles (bounding boxes), which are sampled based on the updated weights.

The sampled particles are output as object proposals and they are guided by the image and history updates.

RESULTS

We compile and manually annotate a large-scale dataset of urban images with labels for windows and facades.

To further show how the priors help get better proposals, we apply the prior weights as a scoring function for the exhaustive space of proposals retrieved by EdgeBoxes. The figure above shows how adding the priors improves the recall and thus AP of EdgeBoxes.

Running time in seconds for 3000 proposals.

<table>
<thead>
<tr>
<th>Method</th>
<th>Finding Proposals</th>
<th>Asset Classification</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geodetic</td>
<td>0.18</td>
<td>4.32</td>
<td>4.5</td>
</tr>
<tr>
<td>EdgeBoxes</td>
<td>0.1</td>
<td>1.26</td>
<td>1.36</td>
</tr>
<tr>
<td>Ours Probabilistic</td>
<td>0.09</td>
<td>0.24</td>
<td>0.33</td>
</tr>
<tr>
<td>Ours Adaptive</td>
<td>-</td>
<td>-</td>
<td>1.1</td>
</tr>
</tbody>
</table>

Running time in seconds for 3000 proposals.