Object Detection in the DCT Domain: is Luminance the Solution?

Deguerre Benjamin
Chatelain Clément
Gasso Gilles

benjamin.deguerre@vinci-energies.net
firstname.lastname@insa-rouen.fr

December 9, 2020
Objectives and proposed solution

What we want to do
- Object detection
- Low computational resources
- Low bandwidth usage

How do we address the problem in this article
- Avoid the decompression step in the processing pipeline
- Follow the compression (JPEG) main idea and focus on important features, more specifically the luminance

Figure: Proposed overall processing pipeline
JPEG Compression

JPEG Pipeline

Figure: Full JPEG compression/decompression pipeline
JPEG Compression: Luminance and DCT

(a) Y, Cb and Cr representation of the data

(b) Y DCT representation of the data
Proposed method

RGB

<table>
<thead>
<tr>
<th>Block 1</th>
<th>Block 2</th>
<th>Block 3</th>
<th>Block 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>h, w, /2</td>
<td>h, w, /2</td>
<td>h, w, /2</td>
<td>h, w, /2</td>
</tr>
</tbody>
</table>

Backbone
Proposed method

RGB

Block 1 -> Block 2 -> Block 3 -> Block 4 -> Block 5 -> Block 6

3

w

h, w

/

2 /2 /2 /2 /2

Backbone

<
Proposed method

RGB
Block 1
Block 2
Block 3
Block 4
Block 5
Block 6
DCT

Backbone

Deguerre Benjamin |
Detection on PascalVOC/MS-COCO

<table>
<thead>
<tr>
<th></th>
<th>Pascal VOC</th>
<th>MS-COCO</th>
<th>FPS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Input</td>
<td>Y only</td>
<td>Full Input</td>
</tr>
<tr>
<td><strong>VGG based:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSD300 (RGB)</td>
<td>74.0</td>
<td>-</td>
<td>24.5</td>
</tr>
<tr>
<td>SSD300 DCT (YCbCr)</td>
<td>60.0</td>
<td>59.8</td>
<td>14.3</td>
</tr>
<tr>
<td><strong>ResNet50 based:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSD300-Resnet50 (RGB)</td>
<td>73.1</td>
<td>-</td>
<td><strong>26.8</strong></td>
</tr>
<tr>
<td>SSD300 DCT LC-RFA-Thinner (YCbCr)</td>
<td>67.5</td>
<td>70.2</td>
<td>25.4</td>
</tr>
</tbody>
</table>

**Table:** Detection results on the Pascal VOC 2007 and MS-COCO test sets (resp trained on Pascal VOC 07+12 train/val and MS-COCO train/val).
Conclusion

Main results

- object detection is feasible in compressed JPEG images, with speed improvements while still keeping fairly good accuracy
- luminance seems to be enough for object detection
Thank you for your attention!
(Poster session T1.9)