Offline 1000-Class Classification on a Smartphone

Yoshiyuki Kawano and Keiji Yanai
Department of Informatics,
The University of Electro-Communications,
Tokyo, JAPAN
Background

• Rapid progress of smartphone
  – Obtain enough computational power for image recognition.
  – But, rapid mobile recognition is still a challenging task.

• Large-scale object recognition
  – 1000 class (ex. ILSVRC)
  – More practical system
  – But, smartphone’s memory is limited
Objective

Mobile Recognition System
• recognize many kinds of object
• require no communication with a server
  – Recognition on a smartphone
• Offline recognition system
  – Real-time image recognition
Offline vs. Online

• **Advantage** 😊
  - Able to use anywhere
  - Real-time recognition thanks to *no communication* overhead
  - Not required the server

• **Disadvantage** 😞
  - Computational power and memory size
  - Electricity consumption

© 2014 UEC Tokyo.
Related Work

- **Google Goggles**
  - Specific object Rec
  - Similar image search
  - OCR

- **Leaf snap (Kumar et al, ECCV’12)**
  - Identifying *plant species*
Image Recognition

• Image Features:
  – Color Patch FV (only mean)
    • 24 dim local color descriptor
  – RootHOG Patch FV (only mean)
    • 32 dim local RootHOG descriptor
    • Similar to RootSIFT
  – SPM
    • Level 1 (1x1+2x2)
  – FeatureSize
    • Color FV: 7680dim, RootHOG FV: 10240dim
Classifier

- Linear Classifier:
  - AROW
    - Online classifier
    - one-vs-rest

Independent of the number of samples

computation: $O(N)$

memory: $O(N)$

- Weight vectors
  - compressed by scalar-based quantization
Recognition Step

We use light weight features and quantized weight.

Processing over 4 cores in parallel

Takes only 0.27 seconds
Performance

• Top-5 classification rate on ILSVRC

<table>
<thead>
<tr>
<th></th>
<th>uncompressed</th>
<th>compressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate</td>
<td>48.7%</td>
<td>47.9%</td>
</tr>
<tr>
<td>Memory</td>
<td>71.7MB</td>
<td>9.0MB</td>
</tr>
</tbody>
</table>

• Only slight performance loss
  – About 1%
Implementation

- 4 core processing
  - Extract descriptors
  - Feature coding
  - Classifies
- Offline processing
  - Recognition time: 0.27 second
  - 1.6GHz Quad Core (Galaxy Note 2)
Mouse -> cup -> desktop -> web
siamang, gollira, chimpanzee
siamang -> gollira -> chimpanzee