Efficient Object Detection on GPUs using MB-LBP features and Random Forests

Shalini Gupta, Nvidia
Problem overview

- Accurate and real-time object (face) detection on the GPU
Applications

- Smart photography
- Human-computer interaction
Windowed approach

Object/non-object pattern classifier

Most popular algorithms:
Existing solution - Features

- Multi-block Local Binary Pattern (MB-LBP) features

![Diagram showing MB-LBP features](image)

- Multi-block Local Binary Pattern (MB-LBP) features

- Formula:

\[
\text{MB-LBP code} = \frac{1}{w\times h} \sum_{i=0}^{w-1} \sum_{j=0}^{h-1} \left( x_i \times y_j \right) + 2
\]

- Example:

<table>
<thead>
<tr>
<th>B0</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B5</th>
<th>B6</th>
<th>B7</th>
<th>B8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

- MB-LBP code: 01011010
Existing solution - Classifier

- Adaptive boosting cascaded classifier

Sub-window

Stage 1

Stage 2

Stage 3

Rejected sub-windows

More stages ~15-20

Only 3x speedup on GPU
Proposed solution

- MB-LBP features + Random Forest Classifier

Independent decision trees that vote
Analogous to a committee of decision makers that don’t talk to each other.
Why random forests?

- Well suited for GPUs
  - Massively parallel
  - Same amount of computation for each pixel
- Previous work
  - Face detection with HAAR features (Belle 2008)
  - Face Recognition (Belle 2008, Ghosal 2009)
  - Expression recognition (Fanelli et al., 2012)
- Fast training
- Possible to add recognition on top of detection
- Online learning
Random forest training

- Train multiple independent decision trees
  - Each tree trained on a random subset of data selected via bagging

Randomly picked subset of features determine each split

Each feature represents a possible split
Randomly picks features 1, 5 & 6
Feature 1 is better than 5 & 6 so is chosen for the split.
Training data

Positive cases (~47K faces)

20x24 rotated and mirrored near frontal upright faces

Negative cases (~50K non-faces)

Randomly selected from 10K images
Feature Selection

- All 5796 MB-LBP features
  - Slow training
  - Lower accuracy
- Feature selection based on repeatability
  - Rejected features selected < 6 times in ~1K trees
  - 2135 features selected
  - Improved accuracy

![Feature Selection Diagram]
Bootstrapping

Positive Cases → Train → Find false positives → Append

Negative Cases

Up to five stages of bootstrapping improved accuracy.
Classifier Parameters

- Ordered decisions
- Increasing number of features randomly selected for a split
- 32 total trees
- Tree depth of 5
GPU (CUDA) Detector

CPU
- Convert to gray
- Resize
- Integral image

GPU
- MB-LBP features
- RF classifier

CPU
- Non-maxima suppression

>95% of computation
CUDA Kernel

- Shared memory
- Thread block
- Bank conflicts
  - Decision trees in cache
  - Trees stored in BFS order as fixed height full binary trees
  - No execution branching while computing trees
Optimizations

- For large images, skip every other pixel - 30% faster
- Reducing bank conflicts by increased bank size and increased registers
- 16 bit integral instead of 32 bit
- Borders and small images on CPU
- Memcopy and kernel temporal overlap
Non-maxima suppression

Final confidence = \( \text{avg}(\text{confidence}) + \frac{\text{(no. of windows)}}{50} \)
- improves accuracy
Accuracy

Measured on the FDDB dataset - 2845 images containing 5171 faces

Hard cases
## Performance (GK107 vs. core i7 - 3.0 GHz)

### Image size 640 x 480

<table>
<thead>
<tr>
<th></th>
<th>MB-LBP + Random Forest</th>
<th>MB-LBP + Cascaded AdaBoost</th>
<th>Haar + Cascaded AdaBoost (Viola and Jones)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU (i7) single core</td>
<td>471</td>
<td>117</td>
<td>200</td>
</tr>
<tr>
<td>GPU (GK107)</td>
<td>22</td>
<td>42</td>
<td>100</td>
</tr>
<tr>
<td><strong>Speed up</strong></td>
<td><strong>21.4</strong></td>
<td><strong>2.7</strong></td>
<td><strong>2</strong></td>
</tr>
</tbody>
</table>

### Image size 1280 x 960

<table>
<thead>
<tr>
<th></th>
<th>MB-LBP + Random Forest</th>
<th>MB-LBP + Cascaded AdaBoost</th>
<th>Haar + Cascaded AdaBoost (Viola and Jones)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU (i7) single core</td>
<td>1752</td>
<td>526</td>
<td>1250</td>
</tr>
<tr>
<td>GPU (GK107)</td>
<td>95</td>
<td>175</td>
<td>425</td>
</tr>
<tr>
<td><strong>Speed up</strong></td>
<td><strong>18.4</strong></td>
<td><strong>3</strong></td>
<td><strong>3</strong></td>
</tr>
</tbody>
</table>
GPU utilization (GK107)

- 95% global efficiency, 5% overhead of loads from shared
- 99.6% occupancy
- IPC ~3
- Further speedup needs algorithmic changes
Conclusion

- MB-LBP features + random forest classifiers for object detection
- Feature selection technique
- Optimized GPU (CUDA) detector implementation
- Highly portable to GPUs (20x speedup)
Questions?