A Real-time 17-Scale Object Detection Accelerator with Adaptive 2000-Stage Classification in 65nm CMOS

Minkyu Kim¹, Abinash Mohanty¹, Deepak Kadetotad¹, Naveen Suda³, Luning Wei², Pooja Saseendran¹, Xiaofei He², Yu Cao¹, Jae-sun Seo¹

¹Arizona State University, ²Zhejiang University, ³ARM, Inc.
Overview of Object Detection Hardware

These smart applications require real-time processing, high frame rates, and low power.

Prior works

- Significant improvement has recently been made in algorithms [1-2], GPUs, FPGAs [3], and ASICs [4-5]
  - Still lacks sufficient accuracy, energy-efficiency, programmability for real-time systems

→ Special-purpose ASIC for versatile object detection

Programmable Object Detection Accelerator

Object detection algorithm

- We employ the Headhunter model based on rigid templates [1]
  - Integrating a large set of weak boosted classifiers, achieving high-speed object detection
  - Combining multiple HOG/LUV channels
  - Achieving ~state-of-the-art face detection accuracy compared to other works

Features

- Multiple classes (e.g., face, traffic sign) that are programmable
- Many objects (up to 50) in one image with different sizes
  - 17-scale support with 6 down-scaling and 11 up-scaling
- High accuracy comparable to state-of-the-art algorithms
  - AP (avg. precision) 0.81/0.72 in AFW/BTSD datasets

Hardware Architecture & Algorithm Adaptation

- Top-level block diagram and data flow

- Algorithm adaptations for hardware efficiency
  - Configurable parameters (e.g., scales, stride, threshold for detection)
  - Weight re-ordering & adaptive classifier cascading

- Hardware optimization techniques
  - Adaptive pooling, pre-processing for NMS function
  - Parallel computation w/ data re-use for multiple search windows
Chip Measurement Results

- **Voltage scaling**
  - 39.8 fps and 159.5mW at 1.0V
  - 16.2 fps and 22.4mW at 0.6V

- **Precision vs. Recall curve**
  - Up to 0.81 AP for face with AFW dataset
  - Up to 0.72 AP for traffic sign with BTSD dataset

*Average Precision (AP) = area under the PR curve*