Rethinking the Camera Pipeline for Computer Vision
or, Building an Approximate Camera

Mark Buckler, Cornell
Suren Jayasuriya, CMU
Adrian Sampson, Cornell
Mobile vision is a pretty cool idea

- object recognition
- object localization
- image segmentation
- 3D structure reconstruction
- localization & mapping
- optical character recognition
- face recognition
- activity recognition
- human pose estimation

...always on.
...on your smartphone.
...in real time.
ISCA 2016

11:00am-12:00pm
Session 6: Neural Networks 3
- Eyeriss: A Spatial Architecture for Energy-Efficient Dataflow for Convolutional Neural Networks [slides]
- Neurocube: A Programmable Digital Neuromorphic Architecture with High-Density 3D Memory [slides]

10:30am-11:00am
Coffee Break

11:00am-12:00pm
Session 1A: Neural Networks 1
- Cnvlutin: Inefficuential-Neuron-Free Deep Convolutional Neural Network Computing [slides]

4:00pm-5:00pm
Session 4A: Neural Networks 2
- EIE: Efficient Inference Engine on Compressed Deep Neural Network [slides]
This project: a programmable camera pipeline.
Let’s approximate a camera pipeline

Design **approximation** into the camera sensor and the ISP

Show how to **retrain** vision models to work on the cheaper, raw data

Measure energy-accuracy **trade-offs** latent in real-world vision applications
Photodiode
Amplifier
ADC
Demosaicing
Denoising
White Balance
Gamut Mapping
Tone Mapping
JPEG Compression
Vision Application

Photography
mode

Vision
mode

RAW image

approximate RAW image

power gated
5-bit logarithmic

JPEG image

Vision Application
Reversing the pipeline
Sensitivity to ISP stages

Normalized error

ISP pipeline stages

- Original
- Demosaic
- + Gamma compress
- + Denoise
- All off

Models:
- LeNet3
- ResNet20
- ResNet44
- Farneback
- SGBM
- OpenMVG
- RCNN
- OpenFace

Note: Crash when all off.
"True" demosaicing.
Demosaicing
- Denoising
- White Balance
- Gamut Mapping
- Tone Mapping
- JPEG Compression

Subsampling.
Demosaicing → Denoising → White Balance → Gamut Mapping → Tone Mapping → JPEG Compression
Demosaicing

Denoising

White Balance

Gamut Mapping

Tone Mapping

JPEG Compression

replace both with in-sensor trickery!
Sensitivity to ADC quantization

LeNet3  ResNet20  ResNet44  Farneback  SGBM  OpenMVG  RCNN  OpenFace

normalized error

bits (linear quantization)
Sensitivity to ADC quantization

normalized error

LeNet3  ResNet20  ResNet44  Farneback  SGBM  OpenMVG  RCNN  OpenFace

bits (logarithmic quantization)
How much energy can vision mode save?

<table>
<thead>
<tr>
<th>sensor</th>
<th>ISP</th>
<th>vision ASIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>137.1–338.6 mW [LiKamWa]</td>
<td>130–185 mW [ON Semiconductor] 250 mW [Hegarty]</td>
<td>204 mW [TrueNorth] 590 mW [EIE]</td>
</tr>
</tbody>
</table>
How much energy can vision mode save?

- Sensor
- ISP
- Vision ASIC
How much energy can vision mode save?

sensor

ISP

vision ASIC

readout (ADCs)
How much energy can vision mode save?
Unresolved questions

Dynamic feedback loop

New signal processing to *improve* learnability

Incremental cost for incremental scene changes

Data movement between sensor, ISP, and application