Cornell University
Computer Systems Laboratory

Reconfiguring the Imaging Pipeline for Computer Vision

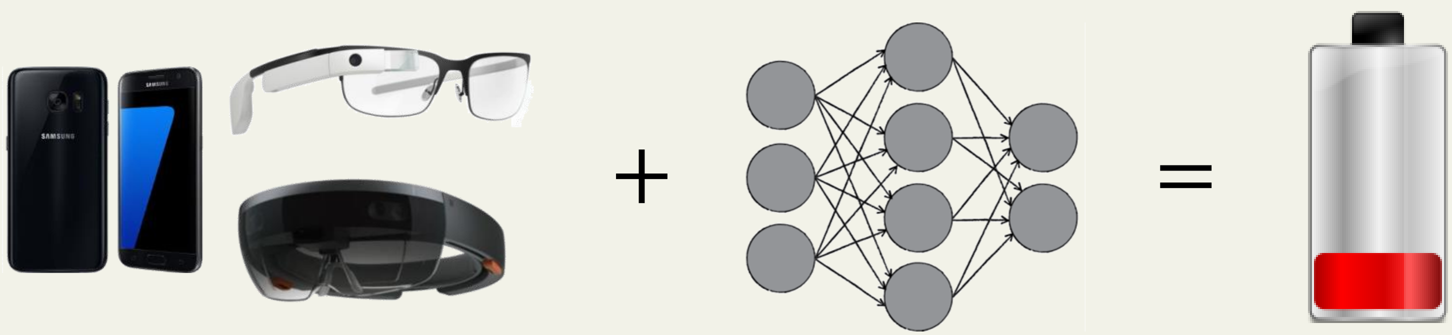
Mark Buckler, Suren Jayasuriya, Adrian Sampson

March 23, 2017

WHERE WE LAST LEFT OFF…

- Deep learning has dramatically increased accuracy for computer vision tasks: face recognition, object detection, etc

- Deep learning and other computer vision applications drain the battery of embedded devices



- Innovation in deep learning ASIC design continues to reduce the cost of embedded inference

- Modifications to the image sensor or ISP have been proposed, but their effect on vision algorithms is unknown

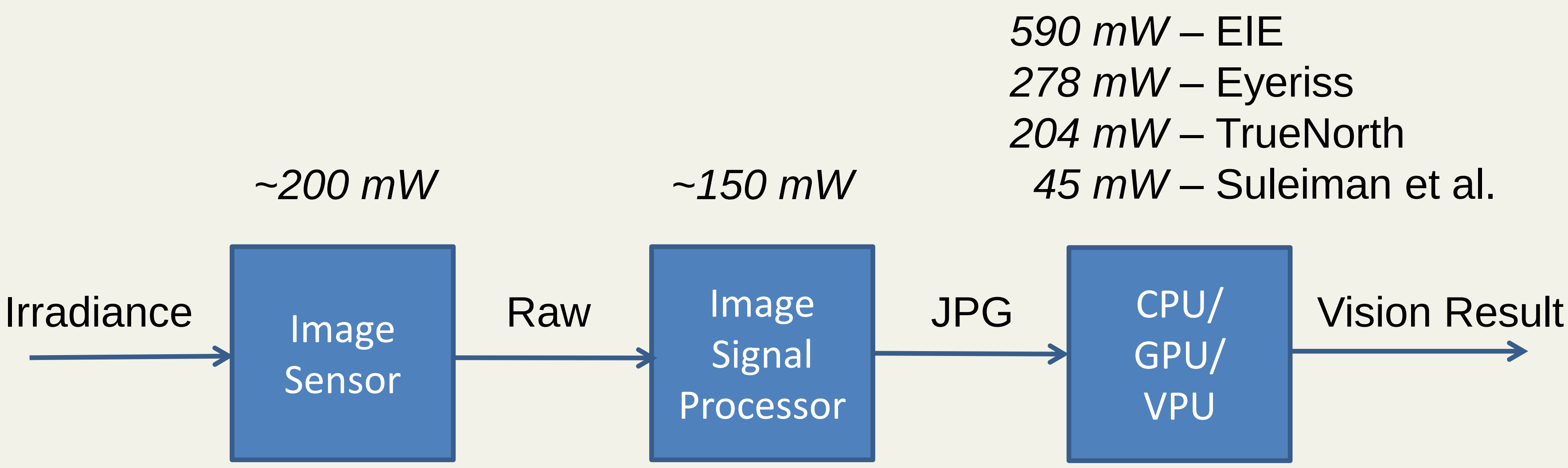
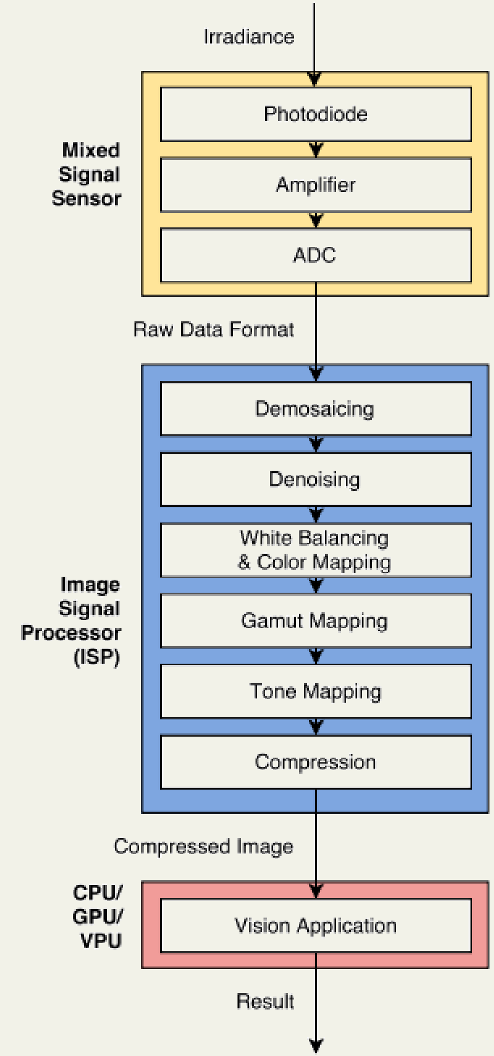


Image Capture for Computer Vision

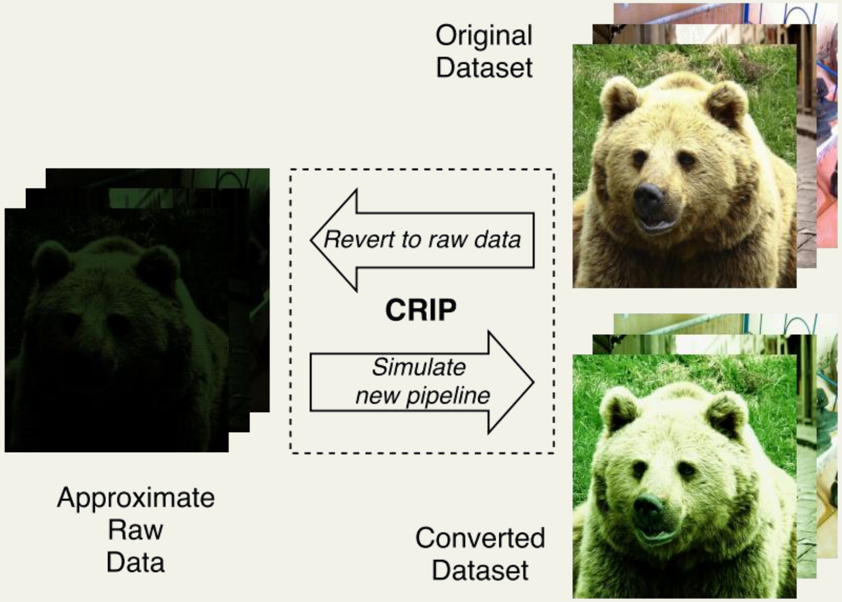
- *Step 1:* Determine computer vision algorithms' sensitivity to sensor approximations and ISP stage removal
- *Step 2:* Use this information to design a configurable pipeline capable of capturing images for both humans and vision algorithms



Evaluating the Impact of Pipeline Changes

- **Nearly all vision datasets consist of human readable images**
- **To train and test vision algorithms on data created by a modified pipeline, we need to convert these datasets**

- ***Configurable & Reversible Imaging Pipeline* (CRIP)**
  - Four stages adapted from Kim et al.'s reversible pipeline
  - Image sensor noise model adapted from Chehdi et al.
  - **Accurate:** <1% error
  - **Fast:** CIFAR-10 in an hour



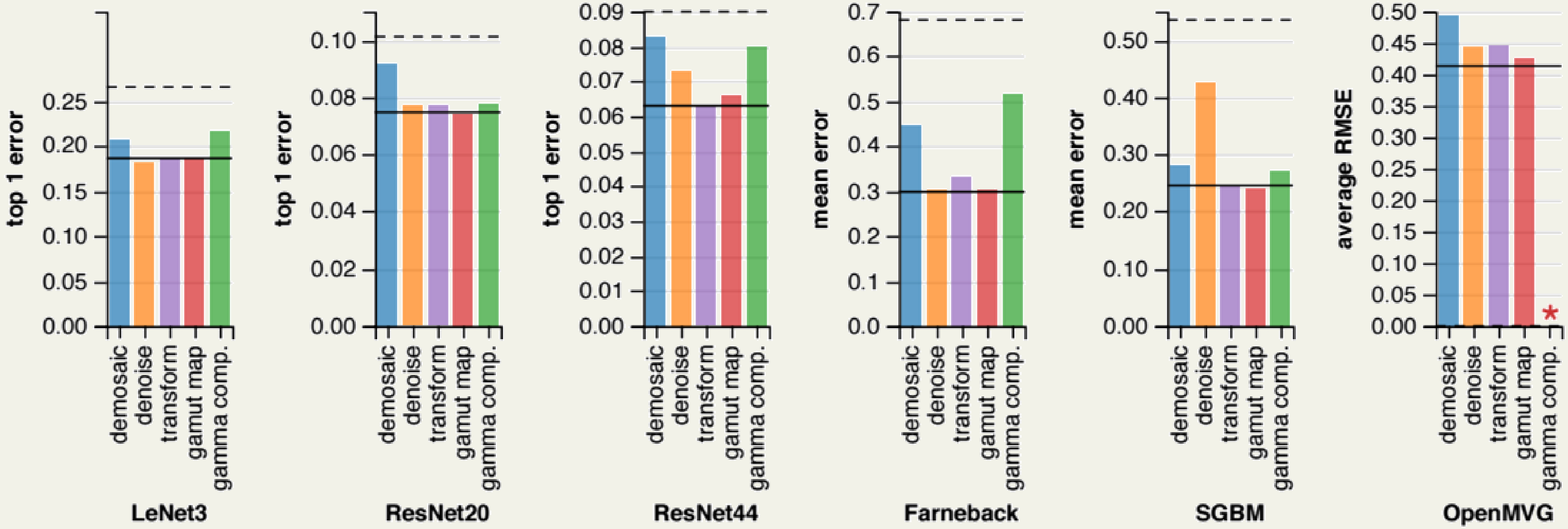
Evaluating the Impact of Pipeline Changes

CSL

- **A wide variety of computer vision algorithms were tested (including deep learning and traditional techniques)**

Algorithm	Dataset	Vision Task
3 Deep LeNet [30]	CIFAR-10 [29]	Obj. Classification
20 Deep ResNet [21]	CIFAR-10	Obj. Classification
44 Deep ResNet [21]	CIFAR-10	Obj. Classification
Faster R-CNN [38]	VOC-2007 [17]	Object Detection
OpenFace [1]	CASIA [46] and LFW [24]	Face Identification
OpenCV Farneback [26]	Middlebury [40]	Optical Flow
OpenCV SGBM [26]	Middlebury	Stereo Matching
OpenMVG SfM [35]	Strecha [42]	Structure from Motion

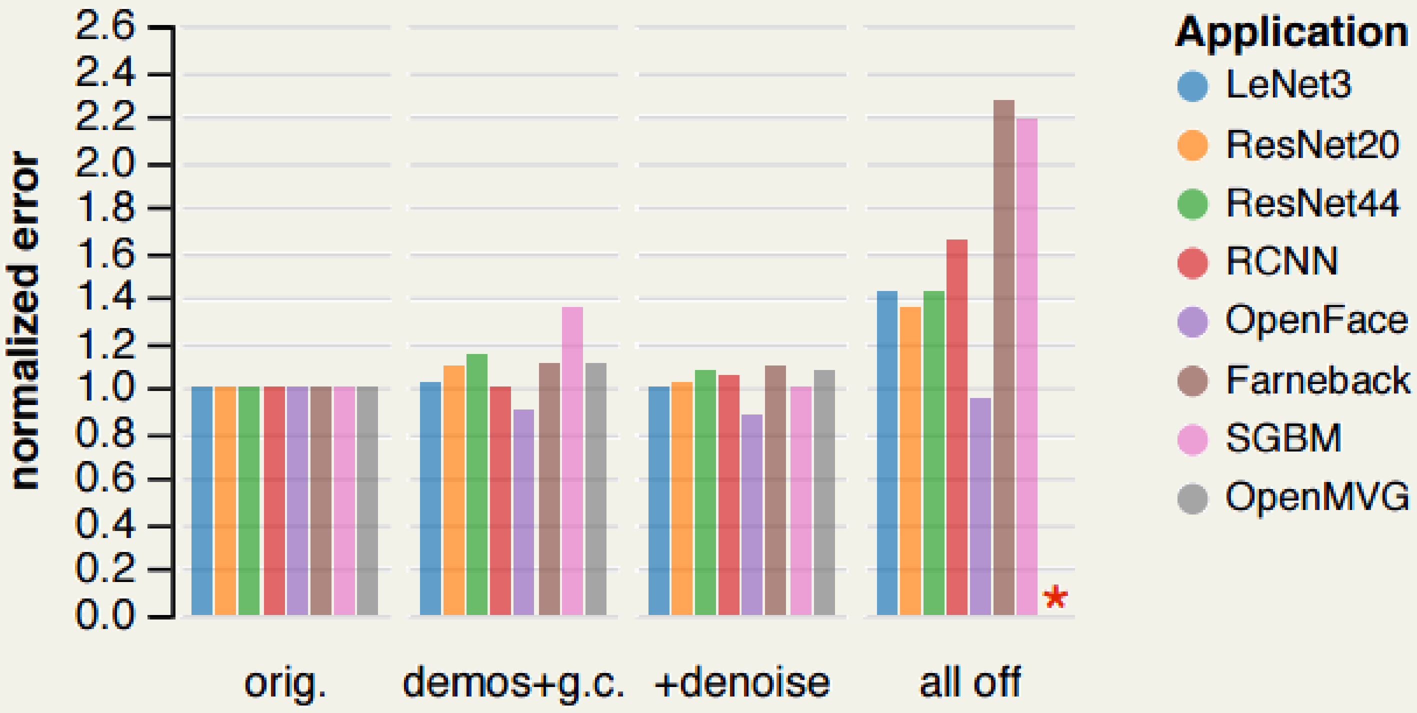
Sensitivity to ISP Stage Removal



Proposed ISP Pipeline

CSL

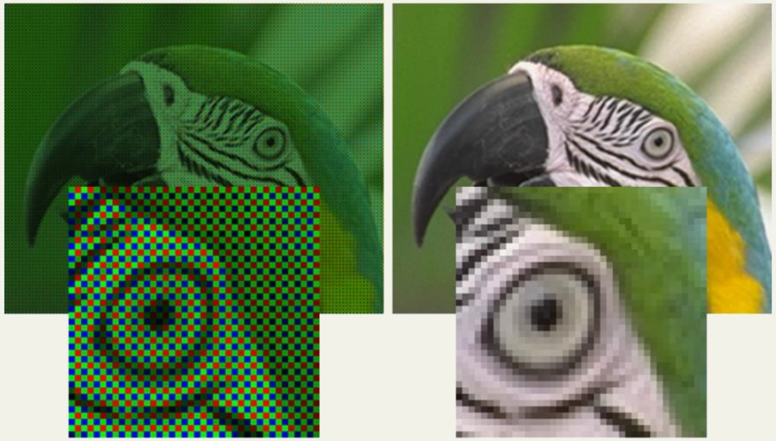
- Most only need demosaicing and gamma compression
- SGBM also needs denoising

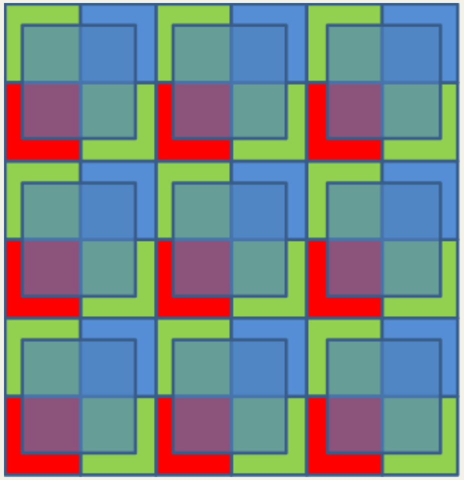


DEMOSAICING: CAN WE APPROXIMATE?

- **Demosiacing algorithms interpolate color values missing from the sensor's filter pattern**

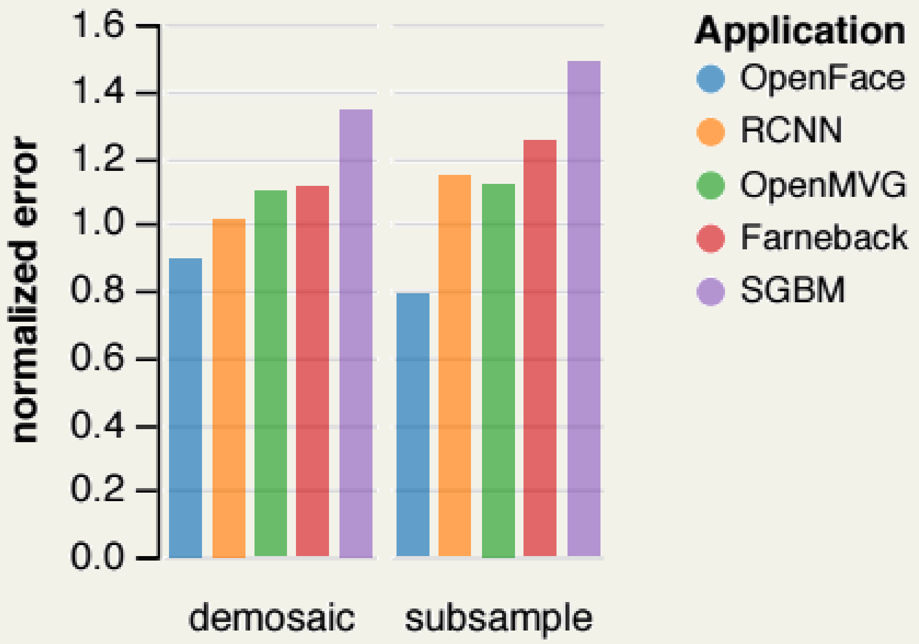
- **Mobile camera resolution >> Network input resolution**
  - Why not subsample instead of demosaicing?





Subsample Demosaicing Results

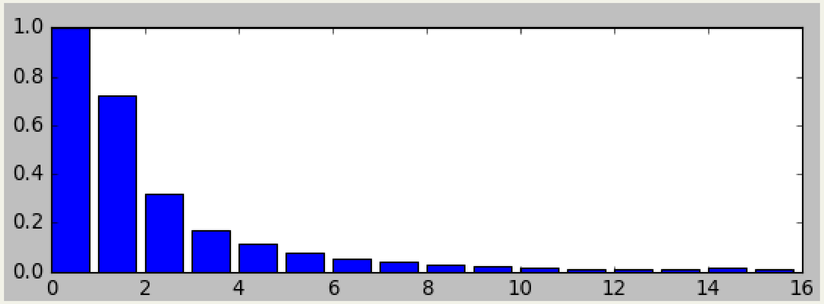
- Tests done with non-CIFAR-10 algorithms
- Tested pipeline contains only gamma compression



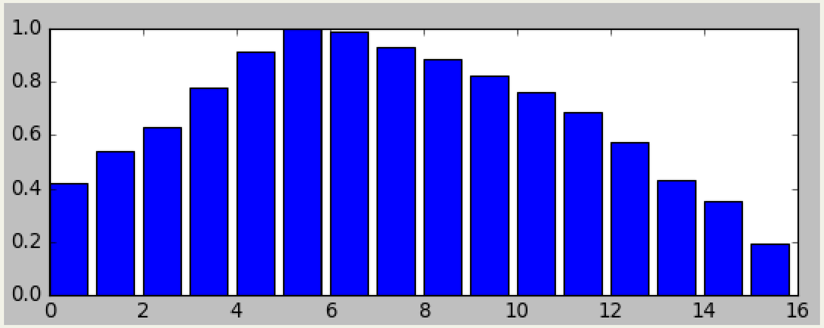
GAMMA COMPRESSION: CAN WE APPROXIMATE?

CSL

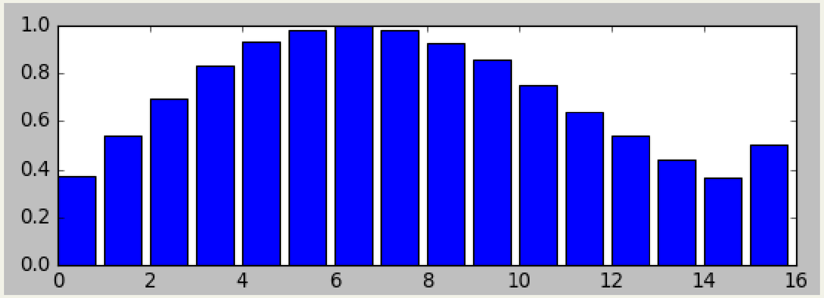
Raw data
(lognormal distribution)



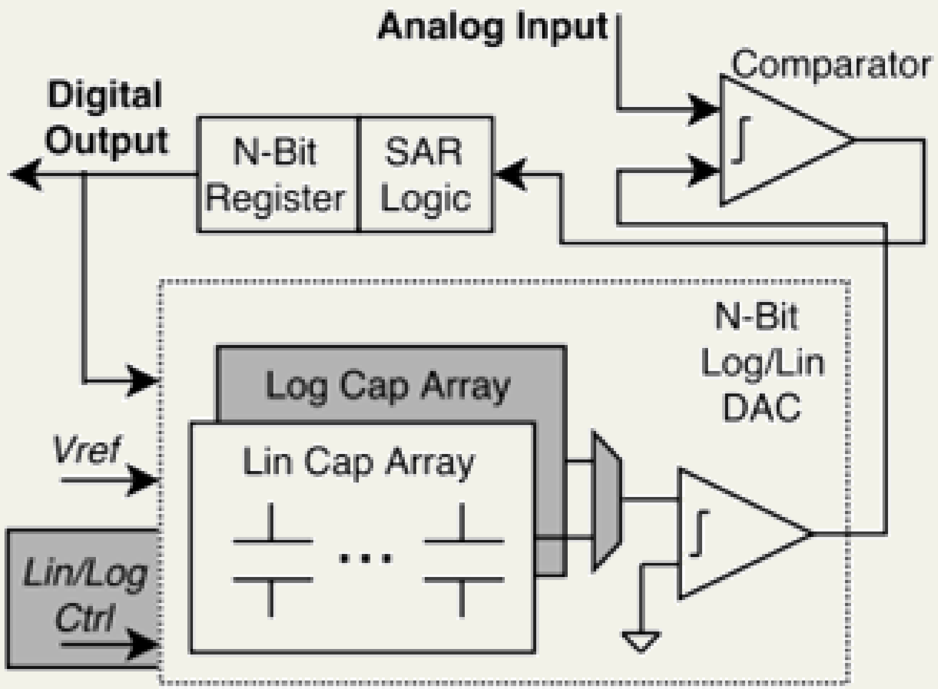
Tone mapped raw data
(normal distribution)



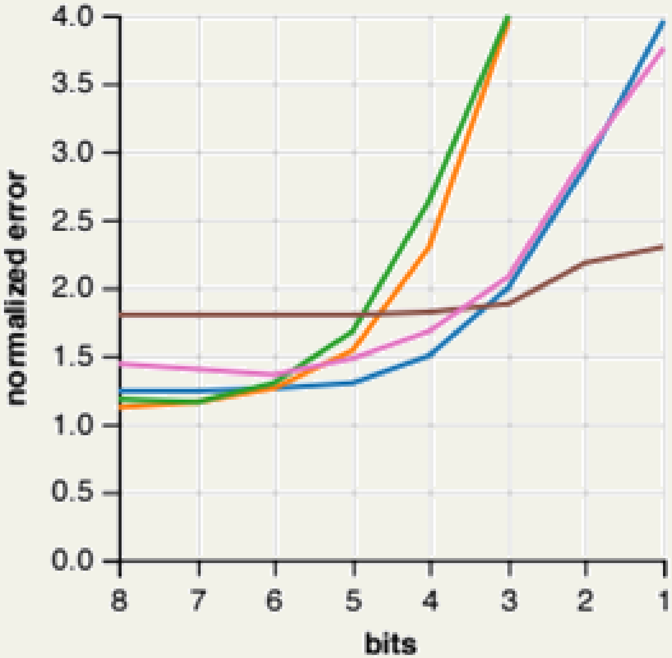
JPEG from standard pipeline
(normal distribution)



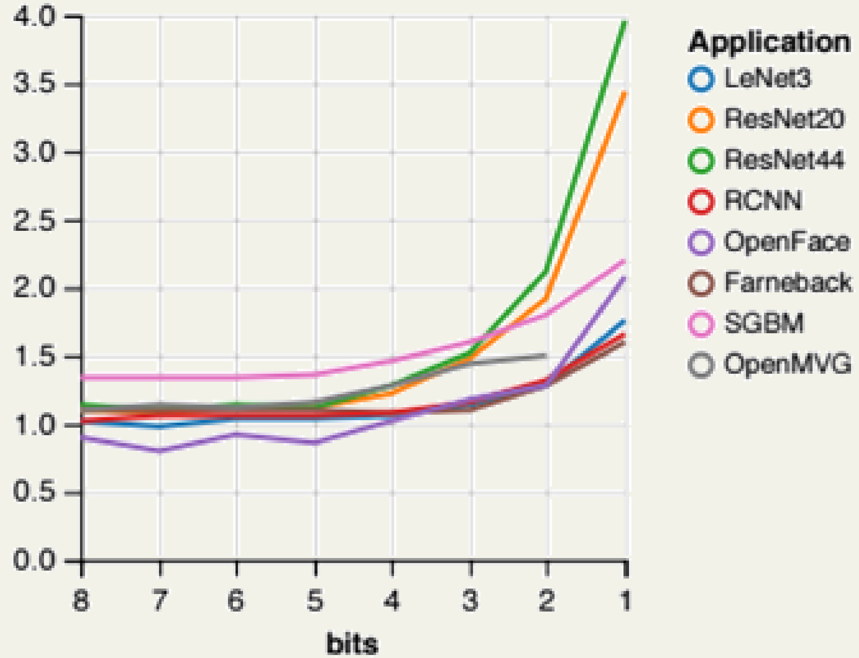
Gamma Compression: Can We Approximate?



Gamma Compression: Use a Log ADC



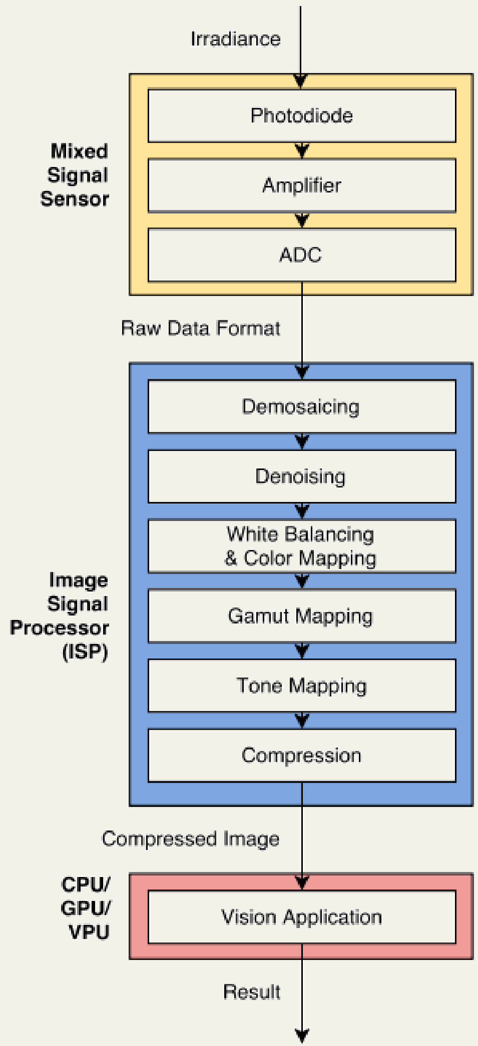
Linear Quantization Sweep



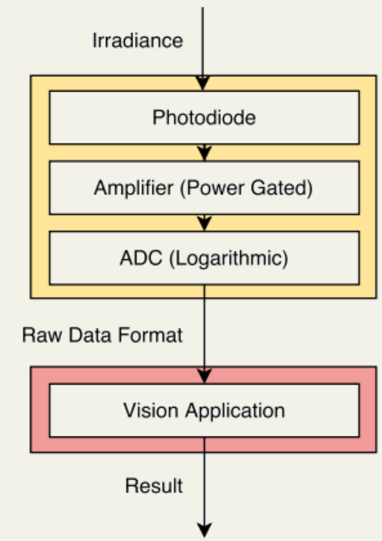
Logarithmic Quantization Sweep

System Design

Photography Mode



Vision Mode



CONCLUSIONS

CSL

1. **All but one application needed only two ISP stages: *demosaicing* and *gamma compression***

2. **Our image sensor can approximate the effects of demosaicing and gamma compression, *eliminating the need for the ISP***

3. **Our image sensor can reduce its bitwidth from 12 to 5 by replacing linear ADC quantization with logarithmic quantization**

Power Savings

- Sensor: ~200 mW, ISP: ~150 mW, VPU: ~300mW
- Half of the sensor energy consumption can be saved by switching from 12 bits to 5 bits
- The entire ISP energy can be saved with power gating

- Our configurable vision mode can save *~40%* of the total system power consumption!