Computer and Machine Vision

Lecture Week 15
Part-1
Wrap-Up Take-away
Outline of Week 15

What Next?
- Future of Computer and Machine Vision
- Merging of CV with Graphics
- Interactive and Prosthetic Vision Systems
- CV Turing Test

Final Review
- Blackboard Online Final Quiz
- Takeaway
Computer Vision

To Go from Machine to Computer Vision, We Need a New Computer and New Algorithms

© Sam Siewert
**MV vs CV vs Video Analytics**

**Machine Vision – Photometers Used in Process Control**
- Successful History
- Industrial Automation and Robotics
- Controlled Environments
- Inspection, Optical Navigation, Medical

**Computer Vision – Emulate Human Vision System**
- Early Underestimation – Marvin Minsky Summer Project
- Challenge of Un-controlled Environments
- 50 Years Later, Challenges Better Understood
- Vision Prosthetics, General Automation
- Recent Breakthroughs – USC, DARPA Artificial Retina, Google Car
- Efficiency and Generalization?

**Video Analytics – CV from RT or Stored/Networked Video**
If Possible, CV => MV Conversion – Cheat!

Practical Solution – Convert CV to MV Problem
- Loss of Generalization (Solves One Problem Rather than Class)
- Controlling Environment May Be Difficult
- Use Non-Visible Spectrum to Advantage (e.g. FLIR)
- Sensor Fusion (Visible + FLIR, RADAR, GPS, …)
- Models and Prior-Knowledge of Problem Exploited

Overhead Camera
Dark Background
Overhead Lighting
Game State / Grid Known
Shape Database
Why is Human Vision > Computer?

1. Neuron > Transistor
2. Better Programming? ROM?
3. More Richly Interconnected
4. Storage + Processing

Red Epic 645
63 Mega-Pixel

Approximately 100+ Mega-Pixel
(Rod/Cone Count)

Cortex=10 Billion Neurons
(High fan-out)

> 1 Trillion Synapses

Total=100 Billion Neurons

5 To 7 billion transistors

CPU

Local Bus

CPU

Memory Controller

I/O Bus (x16 5Gbps = 8GB/sec)

Camera
Link
Interface Card

© Sam Siewert
Biological Vision vs. Machine Vision
(Why A Honey Bee is Better than HPC for CV)

Humans - 100 million Photoreceptors
- 10 billion Neurons (Cerebral Cortex)
- Brain with 100 billion Neurons
- Millisecond Transfer
- Massively Parallel Analog + Digital Computation

Synapse Match is a Challenge
- 7000 Connections from 10 Billion Neurons
- 3 Year Olds Have $10^{15}$ Synapses

CPU to Digital Camera/HDD
- Connects 10’s of millions of pixels
- to Several Billion transistors
- Through Sequential Logic and I/O Bus

960K Neurons in flight: Learns locations, complex odors, colors, and shapes; with high efficiency (500 Watt/Kg), 0.218g


Brain plasticity for learning, connectedness, concurrency, integrated sensing, power efficiency, and resiliency

2016 – 16 billion?

NVIDIA GK110
28nm, (7.1 billion)

Pascal – 15 billion

Intel MICA 22nm
(5 billion)

From Machine to Computer Vision

Problem of:
- Scaling? Efficiency? Latency?
- Architectural Bottleneck?
- Algorithm Challenge?

Storage / Networking Bottleneck
- Access Time in Human Scale is 1 month
- Compared to 1 sec for Memory Access
- Network / Bus is still on Order of an Hour
Inverse Rendering - Computer Vision Turing Test

- HPC Solutions for Digital Cinema (Slower than RT)
- Augmented Reality (RT Interactive)

Scene Description
(E.g. RenderMan, OpenGL)

Scene Analytics

Intelligent Systems

OpenCV and HPC Scaled Processing (ARSC, JANUS)

Rendered Scene

Image


© Sam Siewert
Progress from MV to CV

- Numerous Practical Machine Vision Automation Examples
  - From Optical Navigation to Sorting Recycling Materials, Fish Tagging, Inspection, Agriculture

- Supervisory Control – Robotic Satellite Servicing

- Augmented Reality – Heads Up Information

- Entertainment Systems – Gesture Recognition

- Neural Prosthetics for Computer Vision

- Cameras Designed For Vision (Beyond NTSC, Web and Digital Video)
Neural Prosthetics for Computer Vision

Carver Meade – Neuromorphics
- (VLSI) systems containing electronic analog circuits to mimic neuro-biological architectures

U.S. DoE Artificial Retina Project -
[http://artificialretina.energy.gov/howartificialretinaworks.shtml](http://artificialretina.energy.gov/howartificialretinaworks.shtml)

Argus II – Vision Prosthetic
- [http://www.youtube.com/watch?v=ZyVjK7sktvw](http://www.youtube.com/watch?v=ZyVjK7sktvw)
- [http://www.doheny.org/PDF/Bringing_Sight_to_the_Blind.pdf](http://www.doheny.org/PDF/Bringing_Sight_to_the_Blind.pdf)

Numerous [Additional Projects](http://ine-web.org/about-ine/about-ine/index.html) in Progress
Photometers Designed For Vision
(Beyond NTSC, Web and Digital Video)

- **Eye Camera**

- **Light Field Cameras**

- **3D Scanners**

- **Camera Arrays**
  - [https://graphics.stanford.edu/papers/highspeedarray/](https://graphics.stanford.edu/papers/highspeedarray/)
Open Source Inference

Much of the Future of Computer Vision May Depend on Scene Inference
- Statistical Matching of Key Points and Features
- Incorporation of Context to Recognition
- Chapter 14 in CMV
- Section II, Chapter 6 in “Computer Vision: Models, Learning and Inference”, by Simon Prince

Inference Engines
- Prolog Rule-Based
  - http://www.gprolog.org/
  - http://www.probp.com/
- Bayesian Inference
CV and Image Processing Libraries

- **MATLAB**

- **Mathematica – Image Processing**

- **OpenCV** - http://docs.opencv.org/

How Do We Advance?

- **Fundamental Research – Biological Systems, Physiology, Connectionist Theory and Algorithms** *(UA-A Program in Complex Systems)*

- **Highly Concurrent Perception Processing with Tight Coupling to Photometers Designed for Vision**
  - Low latency, Motion triggered cameras, multi-sensor low resolution and high resolution channels
  - Direct Interface to Concurrent Programmable Parallel Logic and Mixed Analog/Digital Circuitry for Sensory Fusion

- **Research on a Parallel Perception Interface and Range of Concurrency**
  - Analog and Digital Cameras Ranging from Simple Photometers to CCD/CMOS Detectors
  - CPLD, FPGA Concurrent Digital Transformation (Verilog/VHDL -> OpenCL)
  - GPU SIMD Processing for Parallel (CUDA and StreamProc -> OpenCL)
  - Multi-Core Processing with OpenCL

- **Address Challenging Problems that are Cooperative Supervisory Control Between Biology and Machine Vision**
  - Supervisory Control – What Humans Do Best and What Computing Does Best in Collaboration to solve useful problems of Automation – e.g., Robotic Satellite Servicing
  - Human Prosthetics and Hyper-sensory Systems -
Computational Photometer Project

- Cyber-Sensory Perception Interface – Computational Photometer
- Experimental Hybrid Reconfigurable Logic: and Multi-channel Photometer Interfaces with SIMD/Multi-Core Processing and Scalable SSD
- Goal to Build (or Use) Designed for Vision Cameras – Both Analog and Digital
- Integrated Sensor Fusion – E.g. IMU sensors and Cameras
- Leverages and Extends Research at CU-Boulder Funded by Intel – CU-Cam (Analog cameras, CPLD, USB Interface to ARM Coretex)
Human Perception

More than 30 Human Senses Used for Interaction with the World and Self-Monitoring

- [1] Proprioception: Muscle Memory and motion tracking
- [11+] Interoception: Internal Health and Status (homeostatic thermoception, respiration, suffocation, nausea, thirst, cutaneous, GI, esophageal, gagging, fullness, headache …)
- [1] Chronoception: Sense of Time
- [1+] Nociception: Pain receptors
- [1+] Thermoception

Human Sensor Fusion – E.g., Flavor combines taste and smell
CV Related Open Source

Feature Vector Analysis, Pattern Recognition, Classification
- SIFT (Scale Invariant Feature Transform)
  - http://www.cs.ubc.ca/~lowe/keypoints/
- PCA Analysis (related/unrelated dimensions of observed data)
- ICA Analysis (solves cocktail party problem – combined filtering and selective signal enhancement)
  - http://mialab.mrn.org/software/gift/
- Bayesian Logic - http://people.csail.mit.edu/milch/blog/index.html (Monty Hall Problem - Conditional Probability and Baye’s Rule)
  - \[ P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \]
  - Basket-1=5R, 5B balls, Basket-2=3R, 7B; If Basket selected randomly and ball drawn is R, what is Probability ball came from Basket-1?
    - \[ P(R|B1)=1/2, \ P(R|B2)=3/10, \ P(B1)=1/2, \ P(B2)=1/2 \]
    - \[ P(R) = \frac{P(R|B1) \times P(B1)}{P(B1)} + \frac{P(R|B2) \times P(B2)}{P(B2)} = \frac{1/2 \times 1/2}{2/5} + \frac{3/10 \times 1/2}{1/2} = 2/5 \]
    - \[ P(B1|R) = \frac{P(R|B1) \times P(B1)}{P(R)} = \frac{0.5 \times 0.5}{2/5} = 5/8 \]

Prolog
- http://www.probp.com/
- http://www.gprolog.org/

MV/CV Libraries
- http://opencv.org/