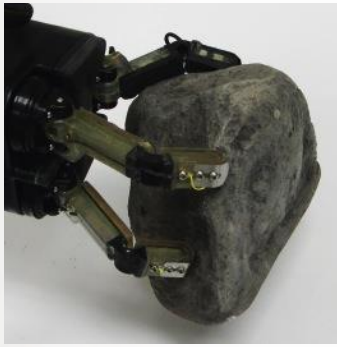
    

Embedding computer vision in everyday life

Mario E. Munich, PhD
VP of Advanced Development

10/2/2013

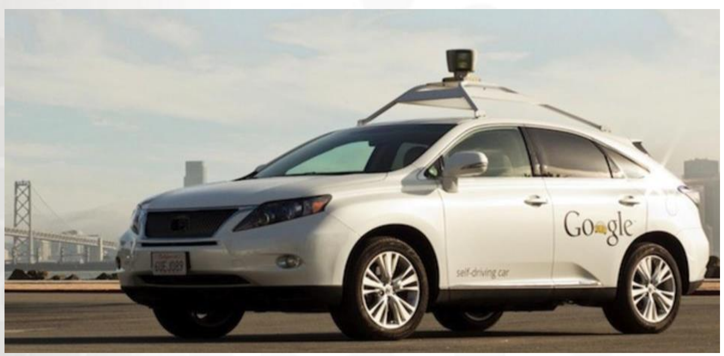
iRobot®
Robots That Make A Difference

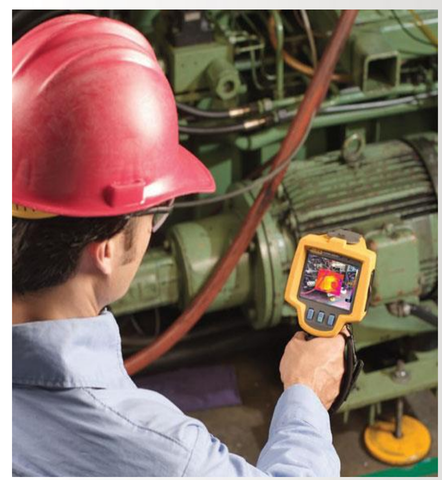
Digital cameras are everywhere

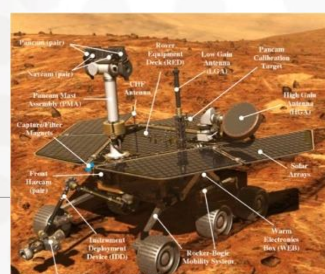










iRobot

Interesting statistics

Digital images are ubiquitous

- Instagram: 27,800 photos/min
- Facebook: 208,300 photos/min
- Flickr: 8 billion pictures uploaded

Projection: 880 billion pictures taken in 2014

Source:PopPhoto.com

Enablers for vision applications

- Affordable components (CPU, memory, cameras, motors, plastics, etc.) to build full systems.
- Powerful, yet portable computational platforms (phones and tablets).
- Pervasive connectivity to vast computational resources ("in the cloud").

Visual Pattern Recognition (ViPR)

Proprietary implementation of David Lowe's SIFT

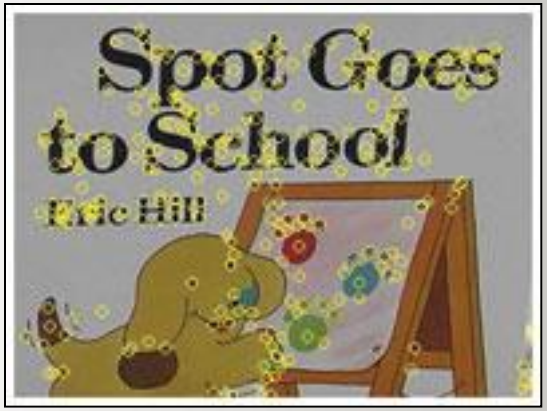
Recognizes individual visual patterns

- Compares images and objects by matching unique features

Supports a variety of applications

- From image database searches to live video analysis
- Scales from servers to embedded systems

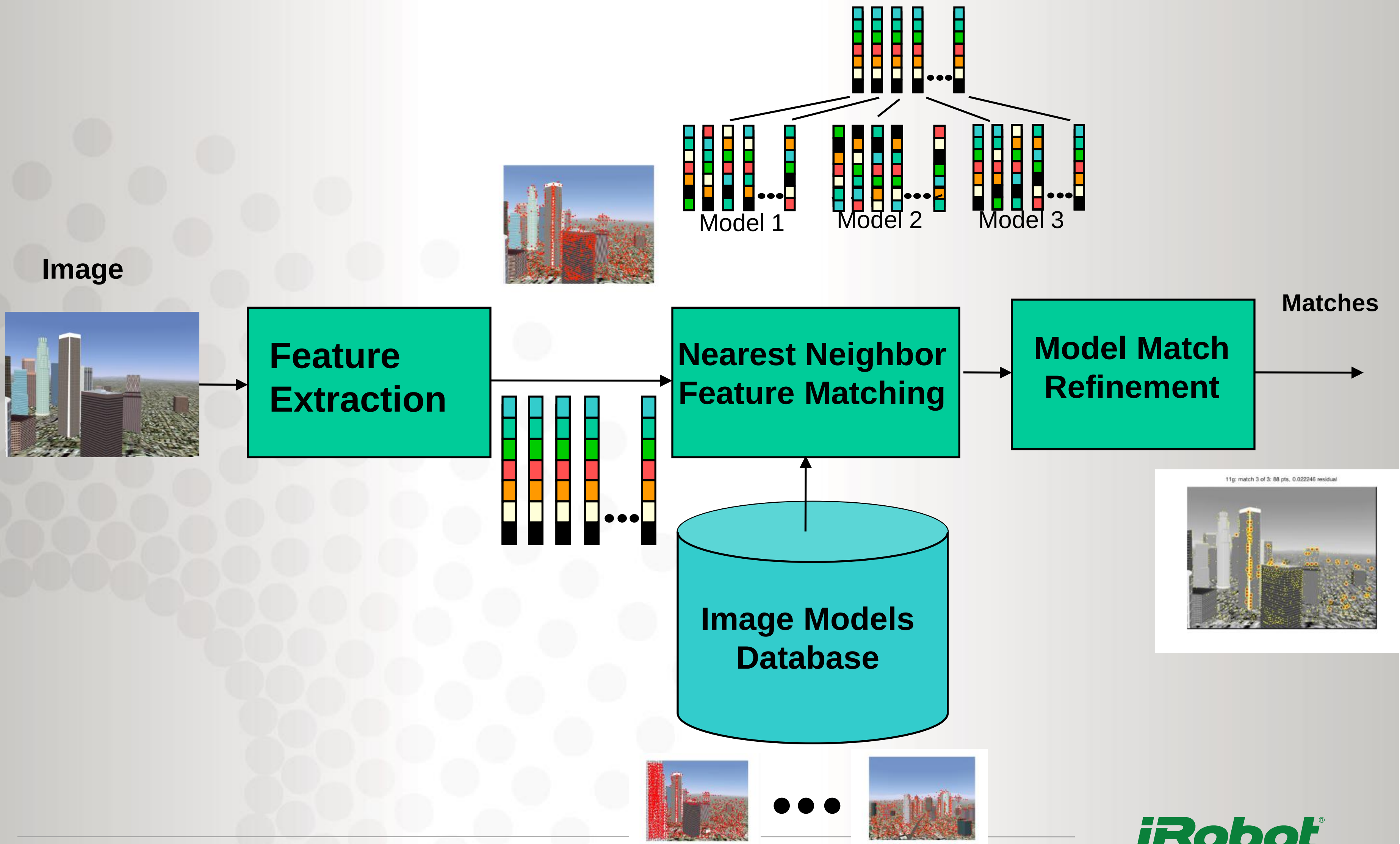
Model



Match

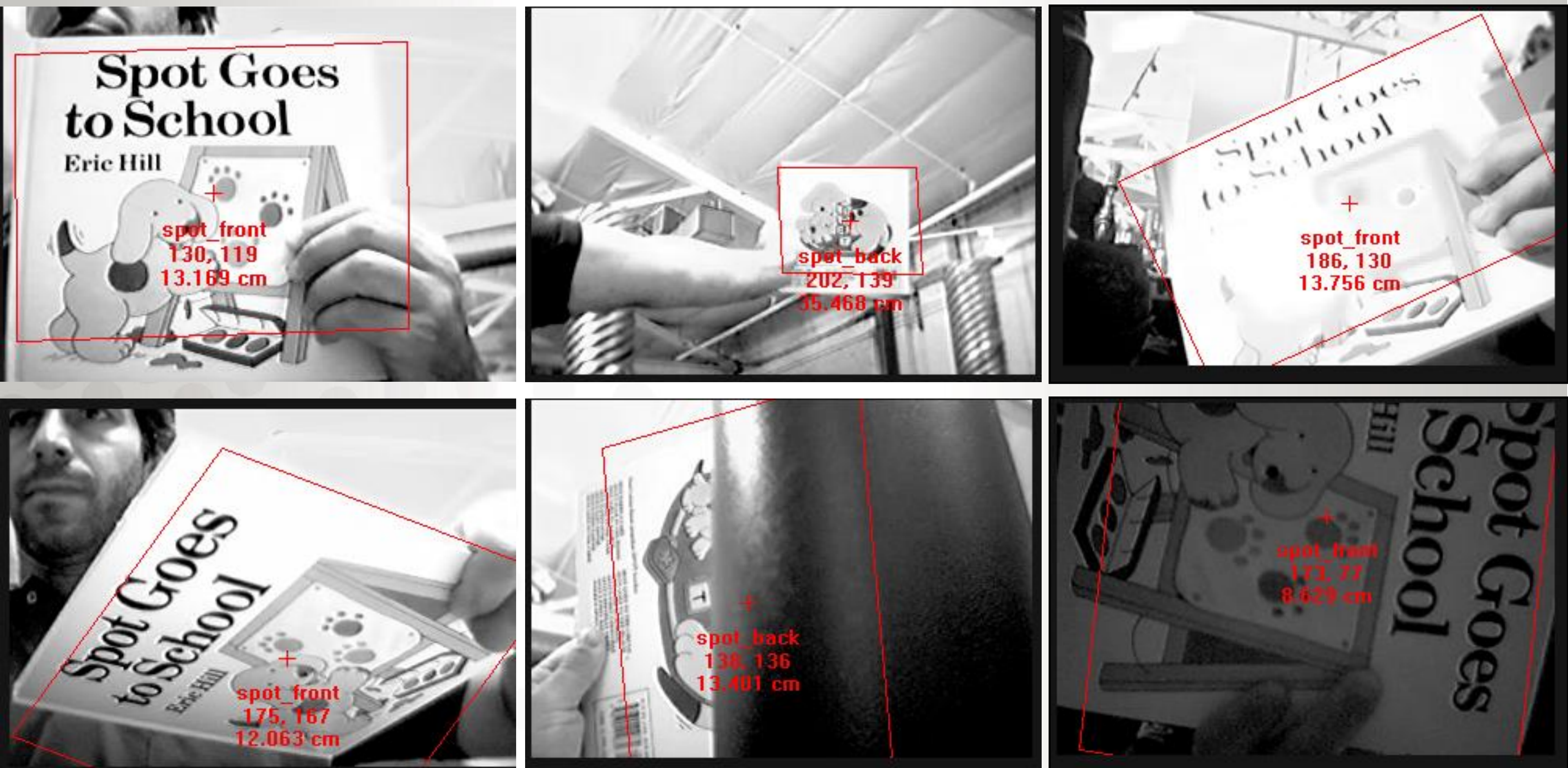
iRobot

ViPR



Model 1
Model 2
Model 3
Image
Feature Extraction
Nearest Neighbor Feature Matching
Model Match Refinement
Matches
11g: match 3 of 3: 88 pts, 0.022246 residual
Image Models Database

Invariance



Rotation
Perspective

Scale
Occlusion

Lighting
Resolution

iRobot

ViPR on Sony's AiBO (ca. 2003)

- Embedded implementation on a MIPS CPU (576MHz)
- Only 30% of the CPU available for ViPR
- Autonomous robot -> ViPR had to work flawlessly in unknown environments

Currency reader for the blind

AD BlackFin BF531 400Mhz built into handheld device
QVGA images

iRobot

Consumer robotics and computer vision

- Cameras can enable smarter capabilities at a lower bill of materials (BOM)
- Challenges of consumer robotics:
  - Tight size, weight, power and cost (SWaP-C) constraints
  - Complicates already difficult integrated SW/HW co-design
    - lens ↔ camera ↔ bus ↔ processor ↔ computer vision ↔ robotic behavior
  - Reliability and robustness

Localization – SLAM with ViPR

What is localization?

- "Where am I?"
- Determine the pose of a robot w.r.t. a reference
- Self-localization, Pose estimation, …

What is it good for?

- Navigation
- Systematic cleaning
- Finding places again

Cost Analysis

Goal: Enable intelligent navigation of consumer robotics products

Rule-of-thumb in consumer electronics/toys:

Production cost: 20% - 30% retail price

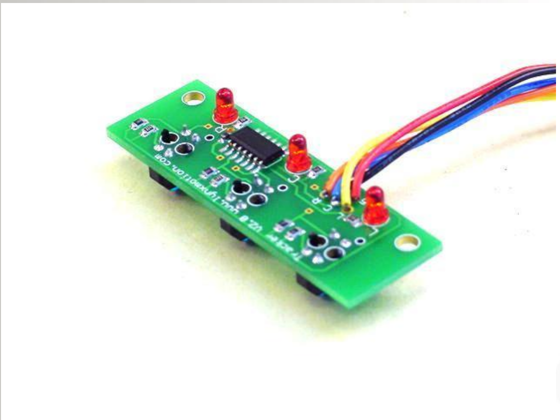
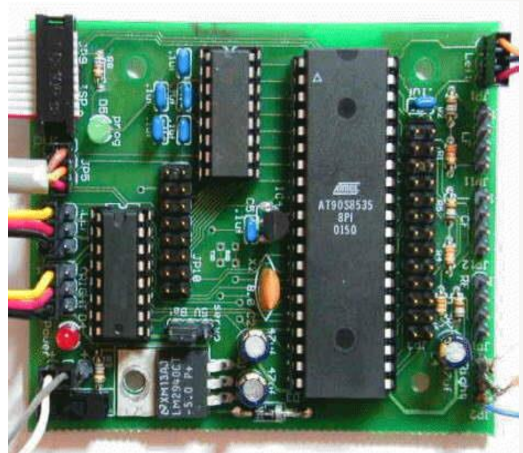
Retail price: **$300**

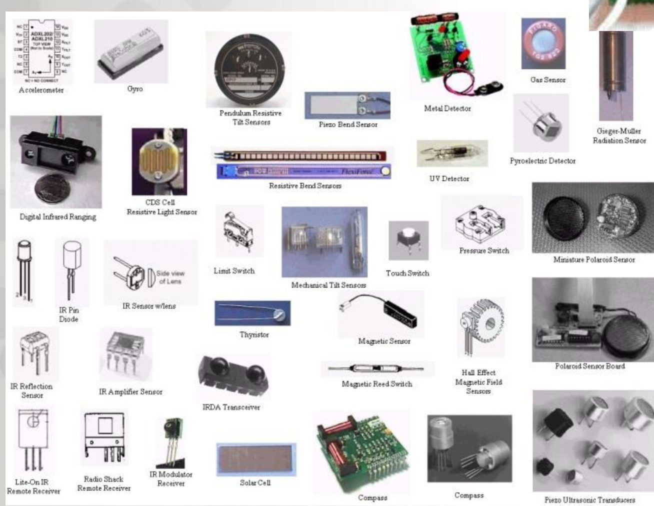
Production cost: **$60 - $100**

(CPU, boards, sensors, actuators, gears, plastics, batteries, charger, vacuum/pads, assembly, etc.)

How much hardware can you get for $60!!!????

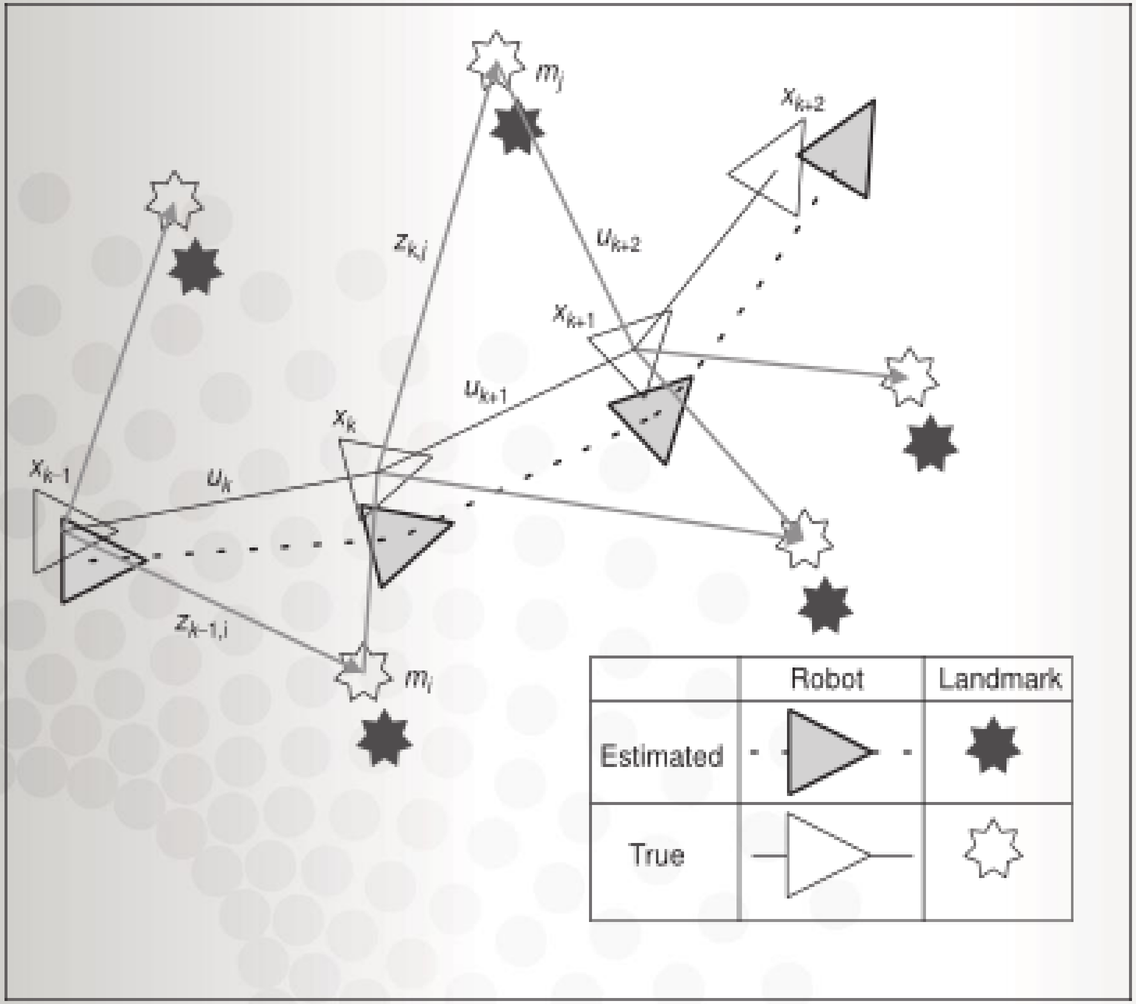
Cost Analysis

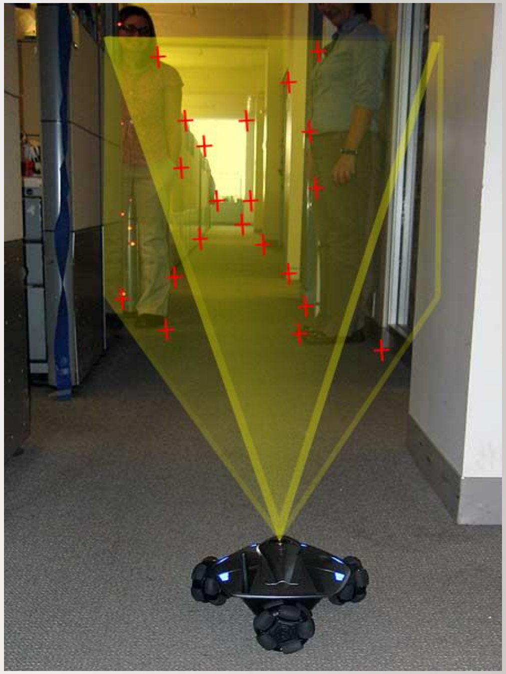



How much robot capabilities can you get for $60!!!????

Visual SLAM (vSLAM)





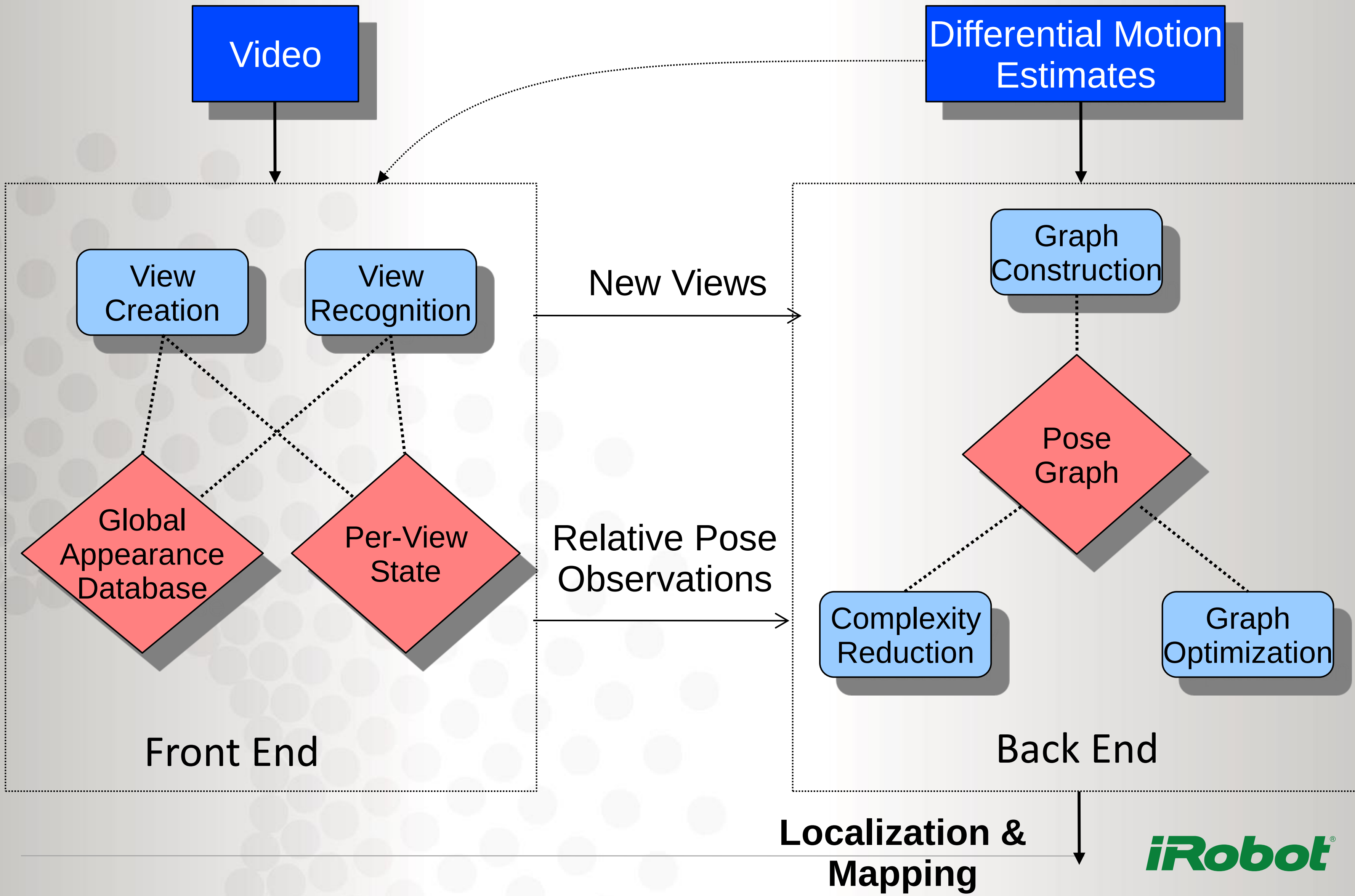
Landmarks:
views of the environment

[Eade, Fong, Munich IROS '10]

SLAM: Map of landmarks AND Pose of the robot

iRobot

System Architecture



iRobot

Views

58 features

74 features

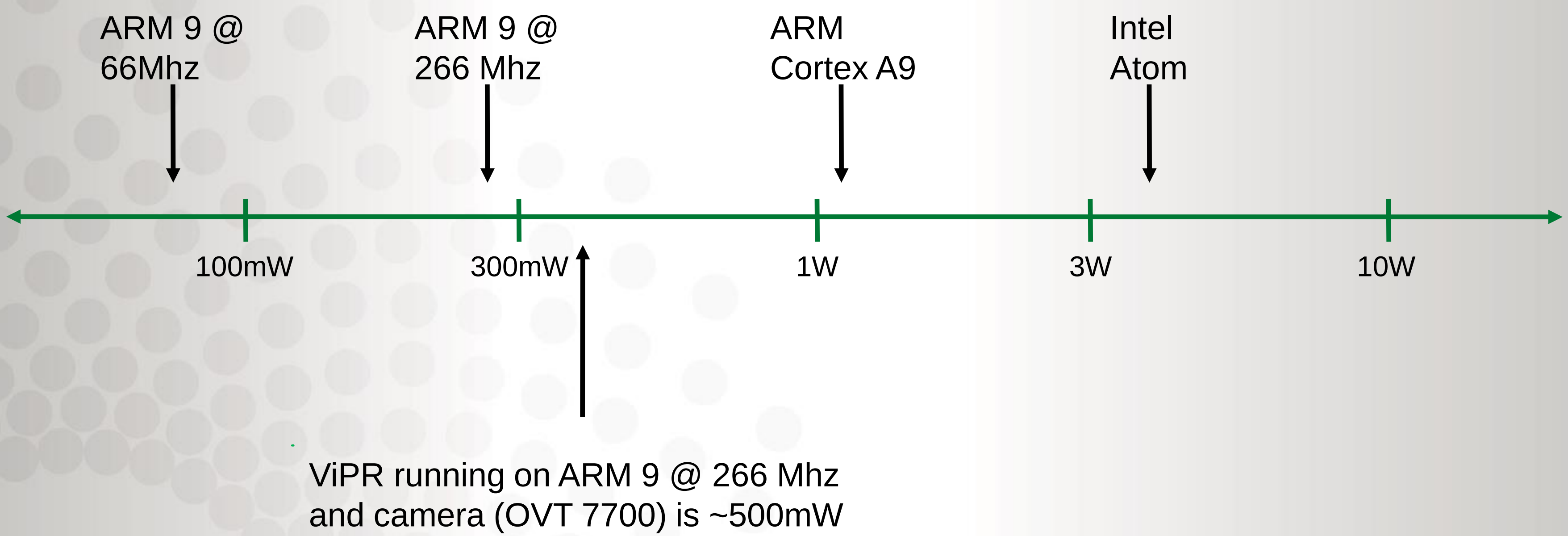
47 features

Front End Timing

Computation profile, ARM9 @ 266 MHz

95 ms	Detect and extract ~200 DoG/SIFT features
40 ms	Choose 3 candidate views
30-40 ms × 3	Perform view-local matching and pose estimation
40-60 ms (if recognition fails)	Attempt to create view
225-315 ms	Front end total

Power Consumption Spectrum



Visual SLAM

evolution robotics™

iRobot®

vSLAM – challenges/trade-offs

- Graph optimization dependent on space and not on time
    - Cost: low tens

- Processor selection:
    - image processing (integer) vs. optimization (floating point)
    - Image sensor interface
    - Clock rate

- Image sensor + lens subsystem:
    - Image size
    - FOV

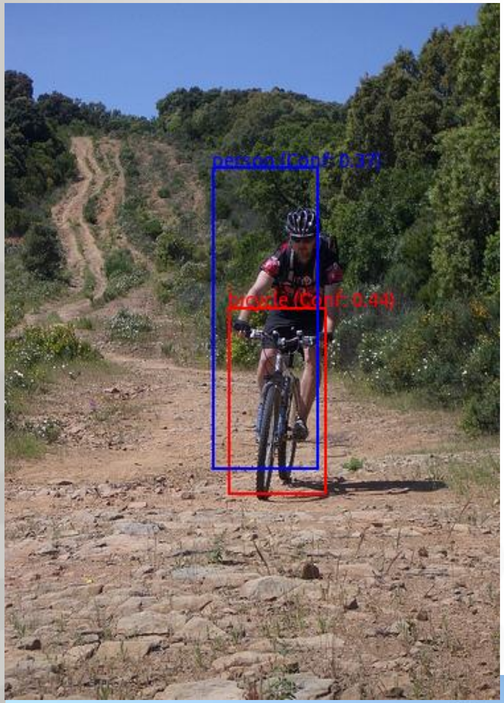
- RAM:
    - cost vs. space to map

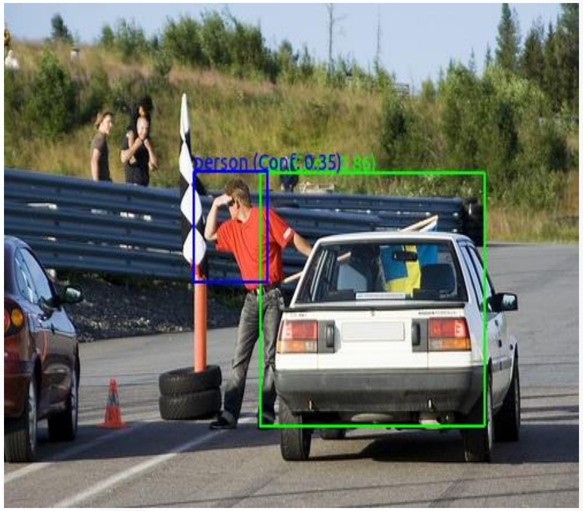
Future: “Human-like Vision”

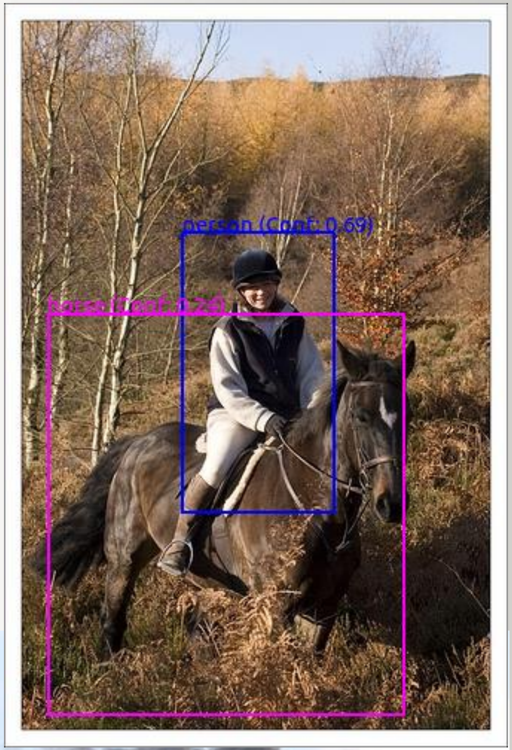
Develop more semantic awareness to enable:

- Smarter navigation capabilities
- More natural and richer human-robot interactions
- Eliminating the need to transmit video from remote robots

iSpot: Category recognition



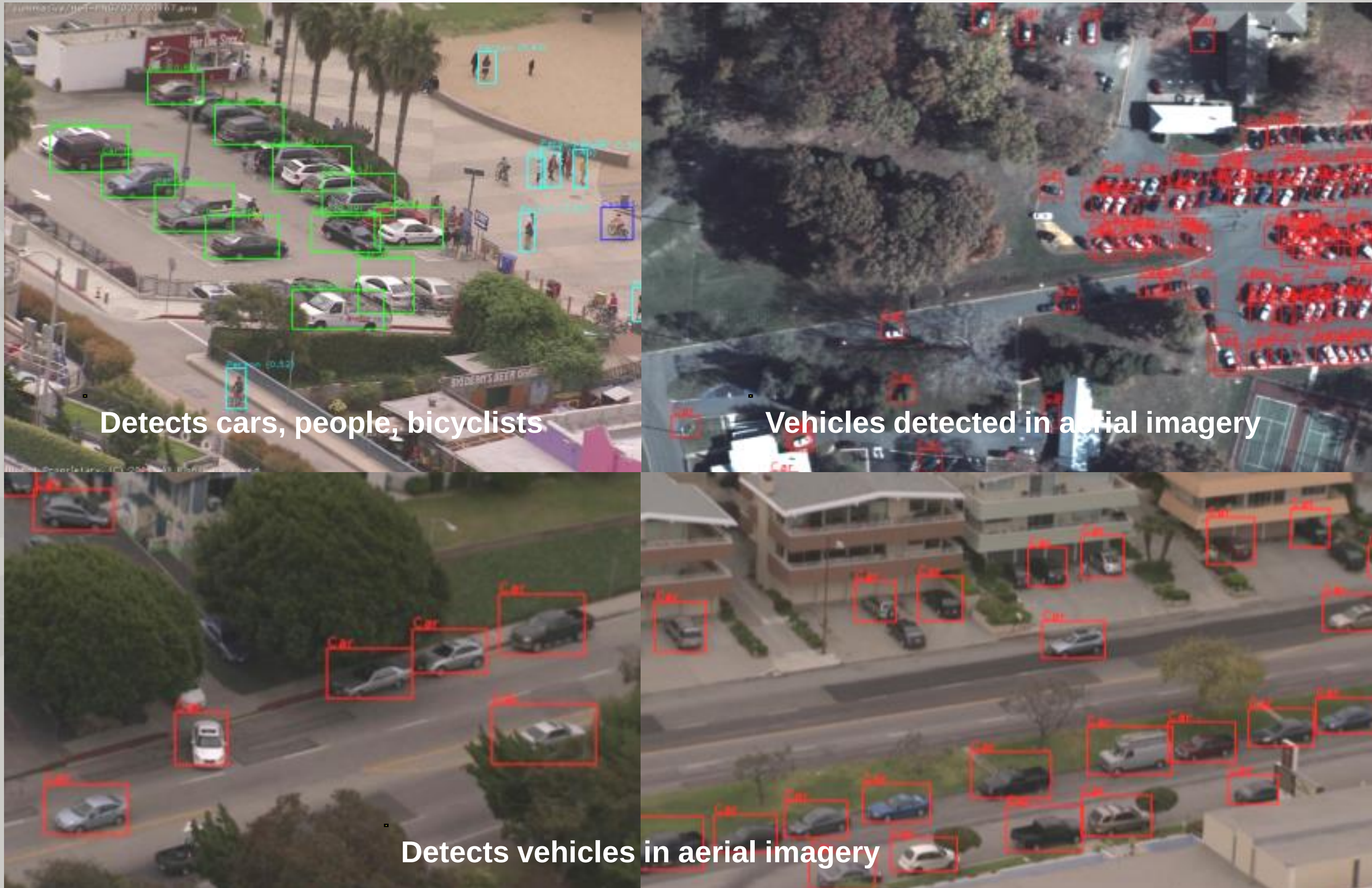




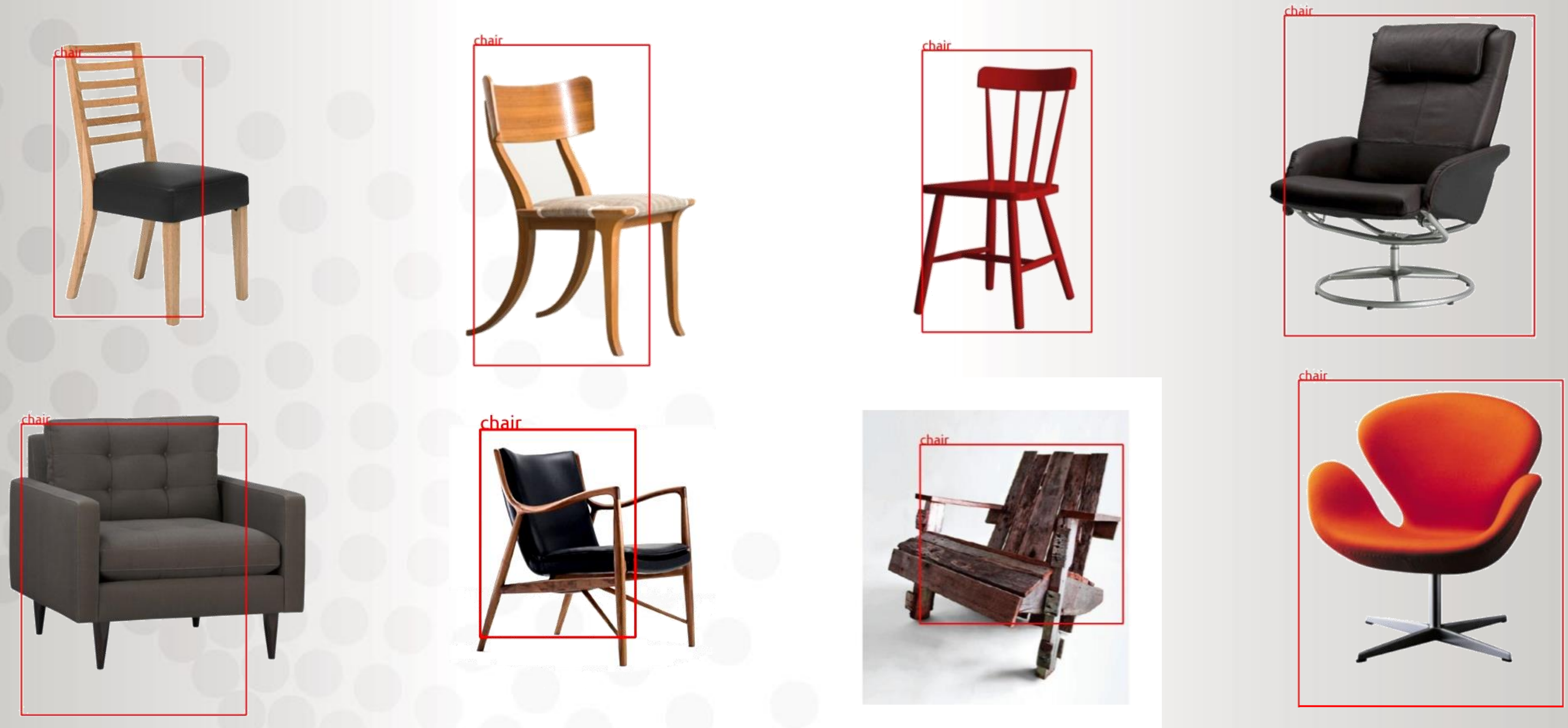
Recognizes people, cars, motorcycles, horses in still photos

iRobot

iSpot: Category recognition

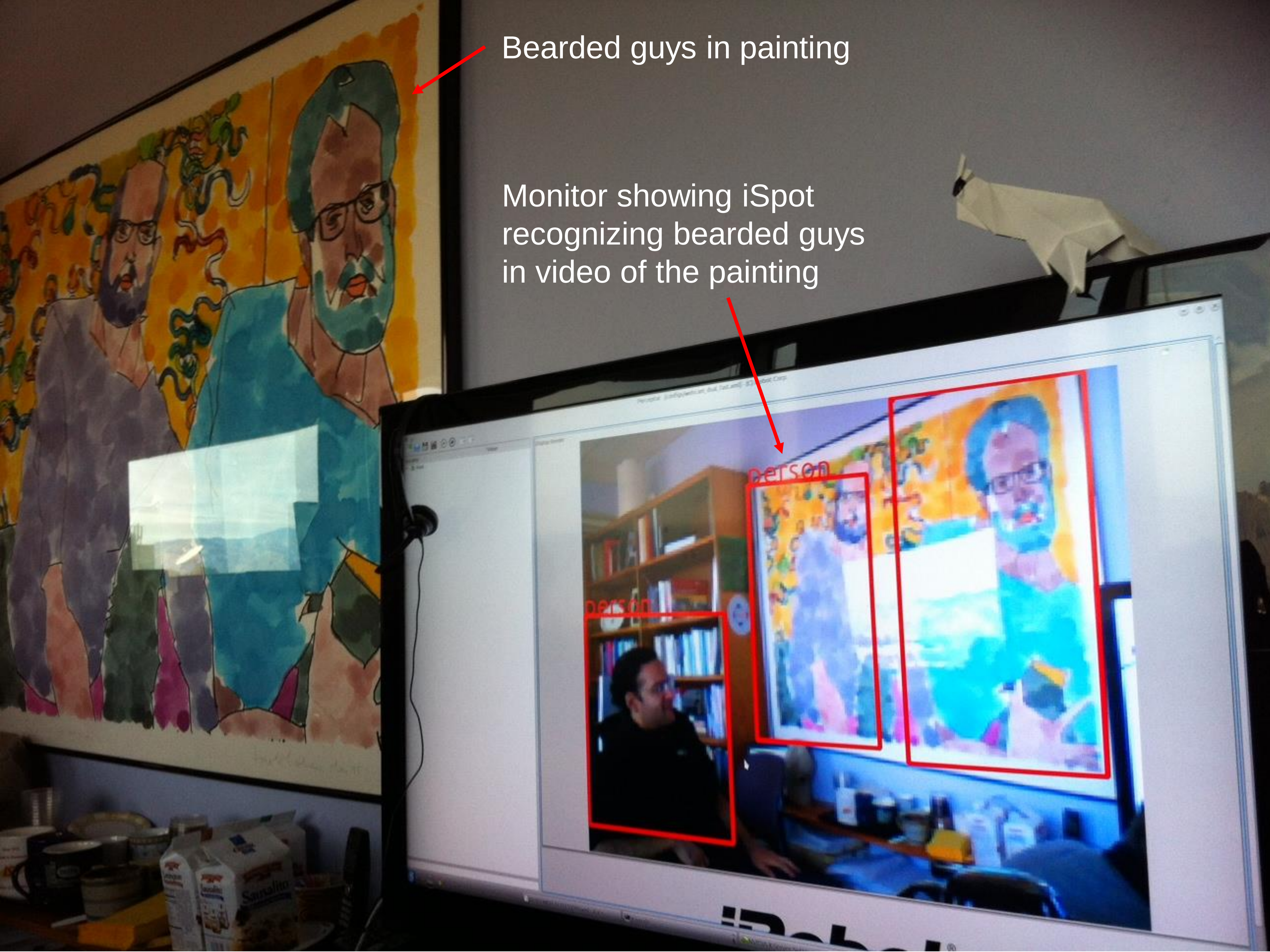


iSpot: Category recognition



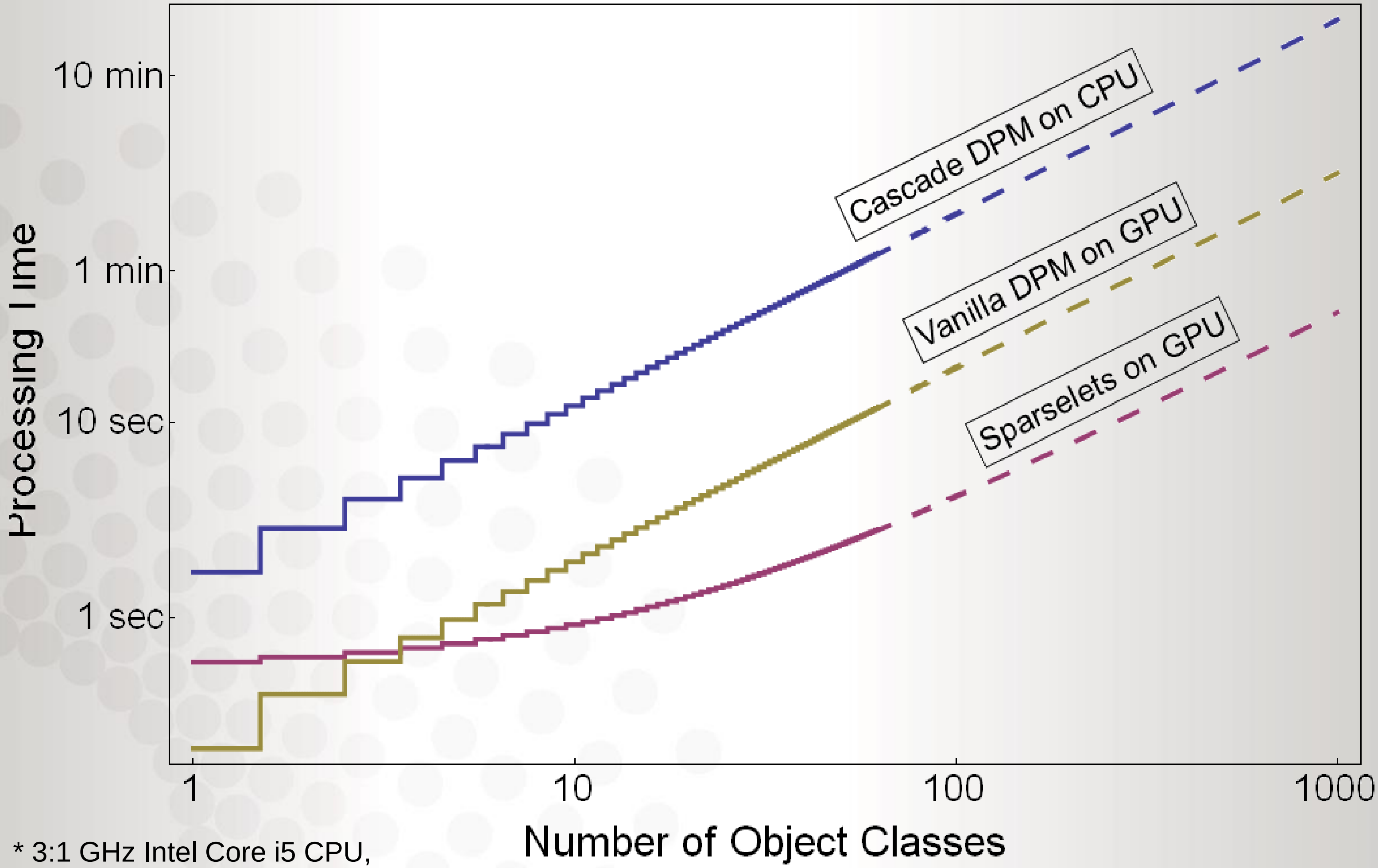

oven (-0.23)
2.89m

iSpot gets categories (even with abstract art)


Bearded guys in painting
Monitor showing iSpot recognizing bearded guys in video of the painting
person
person

iSpot: computation challenge

Deformable Parts Model (DPM) [Felzenszwalb et al., PAMI'09]
Sparselets [Oh et al., ECCV'12]



* 3:1 GHz Intel Core i5 CPU,
Nvidia GTX 580

iSpot – challenges for the future

- Computational requirements -> very challenging for embedding
- Computational requirements -> high power requirements
- Detection vs. False Alarm rate
- Training (100-1000 examples needed)
- Scalability
  - Training
  - Recognition


oven (−0.23)
2.89m

Remote Presence Robots: new platform for vision applications

iRobot

Remote Presence: Ava 500

iRobot


NBCNEWS.COM
#NBCNIGHTLYNEWS

Summary

- Computer vision is an industry in expansion
- The ecosystem (HW/SW) is ready for deployment of applications or services
- Developers can choose to build full embedded systems, deploy on a mobile device or mix devices with the cloud.
- Vision applications will be ubiquitous in the future (mobile devices, robots, appliances, cars, etc.)