Fundamentals of Embedded Computer Vision: Creating Machines That See

Day 1: Introduction to Embedded Vision

September 10, 2012

Jeff Bier
President, BDTI
Founder, Embedded Vision Alliance

What is BDTI?

BDTI is a group of engineers dedicated to helping the electronics industry effectively use embedded digital signal processing technology

BDTI performs hands-on, independent benchmarking and evaluation of chips, tools, algorithms, and other technologies

BDTI helps system designers implement their products through specialized engineering services

BDTI offers a wealth of free information for engineers – visit us at www.BDTI.com

What is Embedded Vision?

- "Embedded vision" refers to embedded systems that extract meaning from visual inputs
  - Embedded vision is distinct from multimedia
- Emerging high-volume embedded vision markets include automotive safety, surveillance, and gaming
  - The Xbox Kinect is the fastest-selling CE device to date: 10 million units in the first 6 months

$130 including game

$920 installed

$300 + $6/month

Why is Embedded Vision Proliferating Now?

1. It has the potential to create huge value
   - Applications in consumer, medical, automotive, entertainment, retail, industrial, aerospace, …

2. It's now possible
   - Sufficiently powerful, low-cost, energy-efficient processors are now emerging

3. Increasingly, it will be expected
   - As embedded vision becomes common in gaming, consumer electronics, and automotive equipment, consumers will expect it

Implementing Embedded Vision is Challenging

- It's a whole-system problem
- There is limited experience in building practical solutions
- Embedded systems are often highly constrained in cost, size, and power consumption
- It's very computationally demanding
  - E.g., a 720p optical flow algorithm, optimized for a modern VLIW DSP architecture, consumed about 200 MHz/frame/second ➔ 5 fps @ 1 GHz
  - Many vision functions will require highly parallel or specialized hardware
  - Algorithms are diverse and dynamic, so fixed-function compute engines are less attractive

Objectives of This Presentation

- Introduce embedded computer vision
- Illustrate representative applications
- Introduce algorithms
- Highlight challenges of implementing computer vision in embedded systems

Applications: Introduction

- Applications of embedded vision are numerous and diverse
- They span almost every major electronic equipment market, including consumer, entertainment, automotive, industrial, security, medical, and aerospace
- In this section we'll briefly look at a few representative applications
- It can be useful to consider the functionality required as distinct from the system and the market
  - Similar functionality may be useful in a variety of systems targeting different markets
  - E.g., gesture-based user interfaces can be useful in smartphones, point-of-sale terminals, industrial equipment, medical devices

Application: Surveillance

- In the U.S., retail theft alone amounts to ~$40 billion per year
- With growing concerns about safety and security, the use of surveillance cameras has exploded in the past 10 years
- The U.K. has led this trend, and has ~1.85 million cameras installed
  - Approximately one camera for every 35 people
  - ~1.85 million cameras generate ~2.5 × 10^9 minutes of video daily
- It's impossible to manually monitor all of this video
- Studies in the U.K. generally show no significant reduction in crime where cameras are installed
- "Smart" surveillance cameras use vision techniques to look for specific kinds of events
- Intelligence can be in the camera, in a local server, or in the cloud
- Key challenge: accuracy with diverse environments and requirements

Cernium Archerfish Solo

Application: Automotive Safety

- ~1.2 million people are killed in vehicle accidents annually
- ~65 million new vehicles are produced annually
- Vision-based safety systems aim to reduce accidents by:
  - Warning when closing in too fast on vehicle ahead
  - Warning of a pedestrian or cyclist in path of vehicle
  - Warning of unintentional lane departure
  - Preventing spoofing of drunk-driving prevention systems
  - Alerting driver when drowsiness impacts attention
  - Automatically dimming high-beams
- Most systems are passive: alert the driver
  - A few apply the brakes
- Some systems augment vision with radar
- Key challenge: accuracy across diverse situations (weather, glare, groups of people, …)

Mobileye C2-270

Application: Video Games

Microsoft Kinect

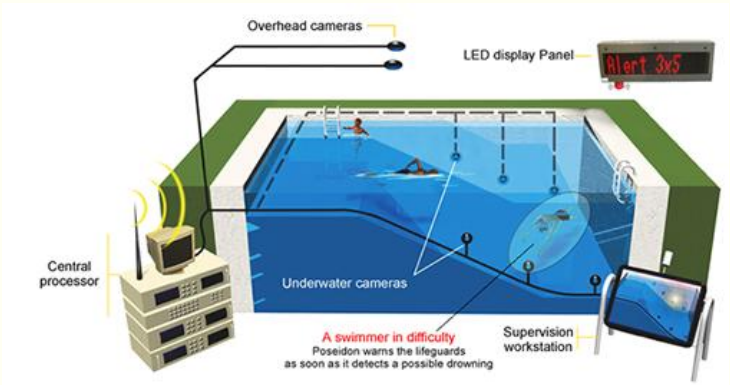
- Video games (hardware and software) are a ~$60 billion/year business
- Vision-based control of video games enables new types of games and new types of users
- Microsoft Kinect: price: $120 (includes a game title); bill of materials cost: ~$60
- Kinect is not just a game controller:
  - Can be used as an audio/video system controller
  - Is being used as a low-cost vision development platform
- Key challenge: must be extremely easy to use and very inexpensive


Image courtesy of and © Useit.com

vs.

Application: Swimming Pool Safety

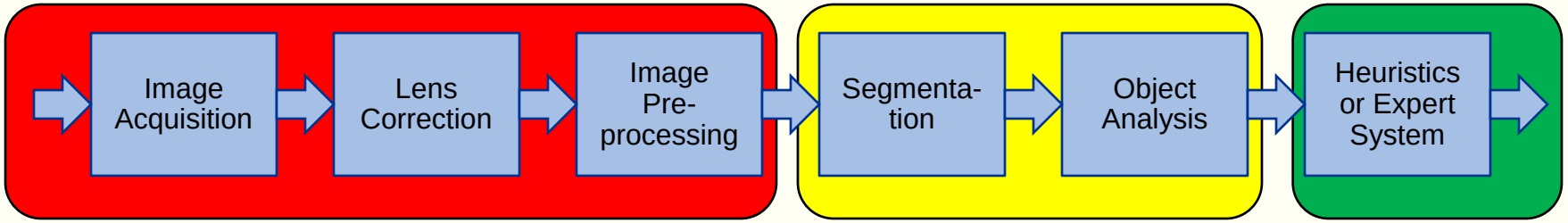
- ~400,000 drowning deaths occur worldwide each year
- In the U.S., drowning is the second-leading cause of accidental death for children 1-14 years old
- 19% of child drowning deaths occur in public pools with certified lifeguards present
- A person drowning is unable to call for help
- The Poseidon system from MG International monitors swimmers and alerts lifeguards to swimmers in distress



Images courtesy of and © MG INTERNATIONAL - POSEIDON

How Does Embedded Vision Work?

- A typical embedded vision pipeline:



Ultra-high data rates; low algorithm complexity

High to medium data rates; medium algorithm complexity

Low data rates; high algorithm complexity

- Typical total compute load for VGA 30 fps processing: ~3 billion DSP instructions/second

- Loads can vary dramatically with pixel rate and algorithm complexity

Lens Distortion Correction—The Problem

- Lenses (especially inexpensive ones) tend to distort images

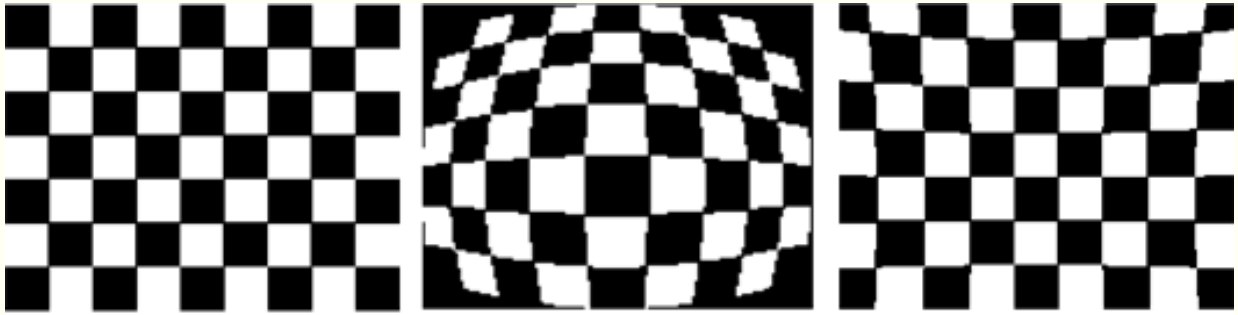


Figure 2. Examples of different types of lens distortion. (a) original (b) barrel distortion (c) pincushion distortion

- Straight lines become curves
- Distorted images tend to thwart vision algorithms

Image courtesy of and © Luis Alvarez

Section based on “Lens Distortion Correction” by Shehrzad Qureshi; used with permission.

Lens Distortion Correction—A Solution

- A typical solution is to use a known test pattern to quantify the lens distortion and generate a set of warping coefficients that enable the distortion to be (approximately) reversed

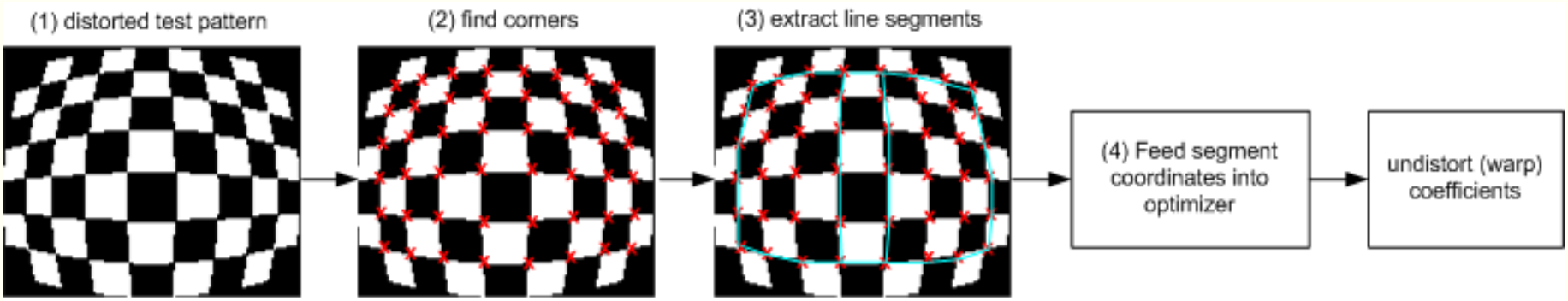


Figure 3. Camera calibration procedure

- The good news: the calibration procedure is performed once
- The bad news: the resulting coefficients then must be used to "undistort" (warp) each frame before further processing
- Warping requires interpolating between pixels

Lens Distortion Correction—Challenges and Trade-offs

- Lens distortion is a well-studied phenomenon, and robust distortion correction solutions exist
  - E.g., http://www.ipol.im/pub/algo/ags_algebraic_lens_distortion_estimation/
- Warping is very computationally intensive
  - Each color component of each pixel requires a calculation
    - E.g., 720p 60 fps: 921,600 pixels × 3 color components × 60 fps ➔ 166 million data elements per second
    - If warping each data element requires 10 math operations (e.g., for bilinear interpolation) ➔ 1.66 GOPS
  - However, warping is readily parallelizable
- There is a trade-off between the quality of the distortion correction and the computation load

Conclusions

To date, embedded computer vision has largely been limited to low-profile applications like surveillance and industrial inspection

Thanks to the emergence of high-performance, low-cost, energy efficient programmable processors, this is changing

In the coming years, embedded vision will profoundly change our industry

Embedded vision technology will rapidly proliferate into many markets, creating opportunities for chip, equipment, algorithm, and services companies – and engineers

But implementing embedded vision applications is challenging, and there is limited know-how in industry

Next Up: Sensors

In tomorrow's presentation, we will introduce sensors, the "eyes" of an embedded vision system.

We'll introduce fundamental sensor characteristics, and explain the operation of the most common kinds of 2D and 3D sensors.

Later in the week, we'll explore:

- Processor options for embedded vision
- Design tools and techniques for embedded vision
- Common building-block algorithms for embedded vision

RESOURCES

The *Embedded Vision Summit*
A Free Educational Event for Engineers—Boston, September 19th

Learn how to use the coolest new technology in the industry to create "machines that see"

- Technical presentations on sensors, processors, tools, and design techniques
- Keynotes by Prof. Rosalind Picard, MIT Media Lab and Gary Bradski, CEO, OpenCV Foundation
- Cool demonstrations and opportunities to meet with leading vision technology suppliers

Part of UBM Electronics' DESIGN East event

- DESIGN East also includes the Embedded Systems Conference, Sensors in Design, DesignMED, Android Summit, LED Summit, and exhibits

The Summit is free, but space is limited. To begin the registration process, send an email to summit@Embedded-Vision.com

For more info: www.embedded-vision.com/embedded-vision-summit

Selected Resources: The Embedded Vision Alliance

The Embedded Vision Alliance is an industry partnership to transform the electronics industry by inspiring and empowering engineers to design systems that see and understand

AMD · ANALOG DEVICES · apical · AVNET · BDTi · CEVA The DSP Powerhouse · CogniMem Technologies, Inc. · cognivue · eyeSight · freescale semiconductor · IHS has acquired IMSresearch · intel · MathWorks · MAXIM · NATIONAL INSTRUMENTS · NVIDIA · omek · PointGrab · SYNOPSYS · TEXAS INSTRUMENTS · TOKYO ELECTRON DEVICE LIMITED · VanGoghImaging A New Vision for Mobile 3D Apps · videantis passion for video · XILINX

Free Resources from the Embedded Vision Alliance

The Embedded Vision Alliance web site, at www.Embedded-Vision.com covers embedded vision applications and technology, including interviews and demonstrations

The Embedded Vision Academy, a free service of the Alliance, offers free in-depth tutorial articles, video "chalk talks," code examples and discussion forums: www.EmbeddedVisionAcademy.com

The Embedded Vision Insights newsletter provides updates on new materials available on the Alliance website. Sign up at www.Embedded-Vision.com/user/register

Embedded vision technology and services companies interested in becoming sponsoring members of the Alliance may contact info@Embedded-Vision.com

Additional Resources

BDTI's web site, www.BDTI.com, provides a variety of free information on processors used in vision applications.

BDTI's free "InsideDSP" email newsletter covers tools, chips, and other technologies for embedded vision and other DSP applications. Sign up at www.BDTI.com.