I. Introduction

II. Smart Camera Architectures
   1. Wireless Smart Camera
   2. Smart Camera for Active Vision

III. Distributed Vision Algorithms
   1. Fusion Mechanisms
   2. Vision Network Algorithms

IV. Requirements and Case Studies

V. Outlook
CHAPTER II: Smart Camera Architectures
Richard Kleihorst, François Berry

Contact: Richard Kleihorst
Richard.Kleihorst <AT> nxp.com
NXP Semiconductor Research

Wireless Camera (WiCa)
Vision Systems

- Are systems that analyze images and video
- They report in events/objects/properties
- DVD recorders, set-top boxes, smart cameras

![Diagram of Vision System]

“Smart Cameras”

- = Camera + intelligence
- = The basis for new applications
  - Such as: detection, tracking, scene analysis

![Applications of Smart Cameras]

- Automotive
- Mobile Comm.
- Surveillance
- Consumer
Some Low-Cost Smart Cameras

CMUcam3 (ARM7)
60 MIPS @ 650mW
Stanford MesyEye Mote (ARM7)

Smart Wireless Camera Platform

WiCa (Xetal SIMD)
50 GOPS @ 600mWatt
Cyclops (AVR RISC)
8 MIPS @ 50mWatt?
Distributed Vision Systems

- Use **multiple cameras** to analyze the scene
- Less problems with occlusion
- Camera **networks**
- Distributed processing
- Distributed reasoning

Requirements in System Integration

- Performance
- Energy consumption
- Cost
- Architectural features (for active vision)
Example Event Casting: Face Detection

Face Detection Application Mapping

Video low level intermediate level high level Data

Pixel processing: Image processing: Application:
Haar filters Image pyramid Draw box, event
for every pixel similar for every image similar For every event different
SIMD 10++GOPS FPGA/DSP 100MOPS CPU 1MOPS

SIMD → Single Instruction Multiple Data
Why SIMD for Low-Level?

- High-performance (need > 10GOPS)
- High internal-bandwidth (need > 500Gb/s)

\[
\text{Bandwidth} = 10\text{GOPS} \times 3 \times 16\text{bits}
\]

Uniprocessor to SIMD: 1 PE

- Performance: 100MOPS
- Size: 5.22mm²
- Performance/area: 19MOPS/mm²
- Overhead: 26%
- Bandwidth: 4.8Gb/s
### Uniprocessor to SIMD: 2PEs

- **Performance**: 200MOPS
- **Size**: 5.24mm²
- **Performance/area**: 38MOPS/mm²
- **Overhead**: 25%
- **Bandwidth**: 9.6Gb/s

### Uniprocessor to SIMD: 100PEs

- **Performance**: 10 GOPS
- **Size**: 8.2 mm²
- **Performance/area**: 1.2 GOPS/mm²
- **Overhead**: 20%
- **Bandwidth**: 480 Gb/sec.
### Uniprocessor to SIMD

<table>
<thead>
<tr>
<th></th>
<th>RISC : 1PE 50MHz</th>
<th>Xetal-II SIMD : 320PE@150MHz</th>
<th>Pentium4 2.4GHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak Performance</td>
<td>0.05 GOPS</td>
<td>100 GOPS</td>
<td>6 GOPS</td>
</tr>
<tr>
<td>Size</td>
<td>6.4 mm²</td>
<td>44.4 (0.18u) mm²</td>
<td>131 mm²</td>
</tr>
<tr>
<td>Performance /area 0.18u</td>
<td>0.008 GOPS/mm²</td>
<td>2.25 GOPS/mm²</td>
<td>0.045 GOPS/mm²</td>
</tr>
<tr>
<td>Overhead</td>
<td>26%</td>
<td>12%</td>
<td>??%</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>2 Gb/S</td>
<td>1.5 Tb/S</td>
<td>58 Gb/S</td>
</tr>
<tr>
<td>Peak Power Consumption</td>
<td>1.0 Watt</td>
<td>59 Watt</td>
<td></td>
</tr>
</tbody>
</table>

### Why is SIMD Low-Power?

- Typical DSP instructions need 4 accesses to memory

```
A

\[ C = A + B; \]
\[ C = A > B \ text{?} \ A : B; \]

B
```
Why is SIMD Low-Power?

- SIMD has multiple PEs in parallel
- Arithmetic always has to be done
- But: Instruction fetch is shared multiple times
- Data (A,B,C) access is shared in multiple-word-wide memories
- Accessing an 8 times wider memory takes half the amount of energy per data entity

SIMD Energy Consumption

- Basis: Convolution
  - Computation
  - Communication
  - Memory access

Without voltage scaling, energy saving levels off
Computational Efficiency Growth (Moore)

[Graph showing computational efficiency growth over time]

Smart Wireless Camera Architecture

[Diagram showing camera architecture: SIMD -> DSP -> CPU]

Event reporting
Connecting The Processors

Feedback loop:
- Frame buffer, working with multiple images
- Look-Up-Table
- Image down-up sizing for pyramid approaches
- Image rescanning for lens-distortion correction

Smart Wireless Camera Platform

“WiCa”

- IC3D/Xetal3 based
- Stereo sensor input
- 50GOPS performance
- Typical 100milli-Watts
- ZigBee node
- Battery powered
- C++ programmed
Smart Wireless Camera PCB

ZigBee module

Battery module

What Have We Mapped to WiCa?

Object recognition applications

CVPR 2007 Short Course
Distributed Vision Processing in Smart Camera Networks
What Have We Mapped to WiCa?

Depth estimation from stereo

What Have We Mapped to WiCa?

Gesture recognition
What Have We Mapped to WiCa?

Face detection: soft edge features

Horizontal soft edges

Vertical soft edges

Which Algorithms Run Easily on WiCa?

• Where much of the application is running on the SIMD
• Where the DSP/CPU is used for limited or occasional tasks only
• Choose appropriate algorithmic basis for scene analysis
  – For example: “feature based”
### Some Power Consumption Results

- Object recognition: 25mWatt
- Face detection: 40mWatt
- Stereo depth estimation: 50mWatt
- Gesture recognition: 15mWatt

### Requirements in System Integration

- Performance
- Energy consumption
- Cost
- **Architectural features (for active vision)**
Outline

I. Introduction

II. Smart Camera Architectures
   1. Wireless Smart Camera
   2. Smart Camera for Active Vision

III. Distributed Vision Algorithms
   1. Fusion Mechanisms
   2. Vision Network Algorithms

IV. Requirements and Case Studies

V. Outlook