On the Design of CMOS Vision Systems on Chip

Angel Rodríguez-Vázquez

IEEE CAS Workshop 2008
Rio de Janeiro, August 27 - 2008
Outline of the Talk

- **Some Basic Concepts:**
  - Concept of Vision Systems and VSoCs
  - Conventional Vision System Architecture
  - Rationale for Using CMOS

- **Architectures for Vision Systems**
  - Progressive Distributed Processing
  - Bio-Inspired Vision System Architectures
  - Shifting the Analog-to-Digital Border

- **Topographical Sensor-Processors**
  - Concept of Concurrent Sensory-Processing
  - Multi-Functional Pixel
  - Rationale for Analog Pre-Processing

- **The Eye-RIS Vision Systems**
  - Architecture
  - The System in Operation
  - Some Prospects for Future Developments
Some Basic Concepts

Concept of Vision Systems

Conventional MV Architectures

Rationale for CMOS Implementation
Concept and Ingredients of Vision

Image Acquisition

Image Processing

Interpretation and Decision

© ANAFOCUS 2008
The Concept of CMOS Vision System on-Chip

All functions for smart imaging and vision systems on chip

Image Sensors → Memory → Data Converters → Digital Signal Processors → Single-Chip Solution

Multi-Function System

© ANAFOCUS 2008
Illustrating the Challenge of Autonomy

Systems capable to make autonomous decisions for vision-based guidance

Image acquisition | Low-pass filtering | thresholding | skeletonization | morphological filtering | signal recognition | road-lanes recognition | actuation on the robot wheels.
Illustrating the Challenge of Speed

Systems to close the Perception-to-Action Loop at Thousands Images-per-Second

Speed ~ 40 m/s

Up to 3,000 frames per second

Laser system

Image acquisition | Low-pass filtering | activity detection | motion estimation | object tracking | Loop control | position prediction | coordinates translation | actuation on the laser system.

© ANAFOCUS 2008
Goals of Artificial Vision Systems

- **Perceive lightness and colour under various illuminations**
- **Detect intensity changes and perform 2-D segmentation**
- **Infer 3-D structures from stereo or motion images**
- **Organize surface and regions into objects of interest**
- **Generate description of objects and recognize them among a potentially large class of objects**
- **Make non-visual inferences about the scene based on visual processing (abstraction)**

Emulate the capabilities of human visual systems

[H.R. Myler, Fundamentals of Machine Vision]
The Vision Processing Chain

<table>
<thead>
<tr>
<th>Task n°</th>
<th>Processing Task</th>
<th>Example</th>
<th>Amount of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Image Acquiring</td>
<td>Light sensing</td>
<td>b × N²</td>
</tr>
<tr>
<td>2</td>
<td>Noise reduction</td>
<td>Lowpass, median filtering...</td>
<td>b × N²</td>
</tr>
<tr>
<td>3</td>
<td>Segmentation</td>
<td>Edge detection</td>
<td>b × N²</td>
</tr>
<tr>
<td>4</td>
<td>Feature extraction</td>
<td>shape detection</td>
<td>N²</td>
</tr>
<tr>
<td>5</td>
<td>Classification</td>
<td>identification</td>
<td>N</td>
</tr>
</tbody>
</table>

N = image width & height
b = number of bits of resolution

람 Huge amount of data at Early Stages

RGB 8-bit coded @ 512 x 512 implies 200Mbit/s data rate

Most data are useless
Illustrating the Computational Demands

Per-Pixel 3x3 Linear Convolution

9 Products
+ 8 Sums

If the image is $N \times M$
If the algorithm contains $N_s$ convs.

17 Operations

At $F$ Frames per second

$17 \times N \times M$ OPS

17 Operations

$17 \times N_s \times N \times M$ OPS

10 convolutions on 8-bit gray-scale VGA images at 100 FPS requires 5 GOPS

Memory operations must be accounted for as well

[Original from Gustavo Liñán]
Conventional Vision System Architecture

- Conventional systems are heterogeneous
  Sensors, Memory, Processors, Communications

- Involve multiple technologies

Data Flow:
- CMOS / CCD sensor: Image Acquisition
- High-performance A/D Converters: Image Coding
- Advanced DSP & High Density Memory: Image Processing and Storage

© ANAFOCUS 2008
Troubles of Conventional MV Architectures

- **Physical separation:**
  - Sensors
  - Processors

- **Heterogeneous technologies**
  - CCD or CMOS for sensing
  - CMOS: FPGA, Power-PC, DSP, etc for processing

- **Bottlenecks at different stages**
  - Image coding
  - Image transmission
  - Image processing

© ANAFOCUS 2008
Photo-transduction Mechanisms

- Incident photons create electron-hole pairs
- The absorption length depends upon the wavelength
- An electrical field separate the hole-electron pairs

Rationale for CMOS

CCD: Hamamatsu S7030
CMOS: Perkin Elmer RL0512T
**Pinned Photodiodes**

**Features:**
- $p^+ n p^-$ structure
- Low-dark current
- Operation similar to photogate
- Decrease reset noise (*Correlated Double Sampling*)
- Better response to blue

During integration phase, photo-generated majority carriers are stored in the depletion region. (Device is fully depleted)
The Concept of CMOS Camera on Chip

Fossum’s Camera on Chip

- On-chip Electronics
- Low-Cost
- Low-Power
- Random Access...

**PIXEL ARRAY**

**READ-OUT MODES**
- Progressive scan
- Window read-out
- Skip read-out (subsampling)

**Timing and Control**

Digital Output

**Column Select Logic**

**Analog Signal Processing**

A/D Converters

Column parallel ADC or serial ADC

- Charge integration
- Gain
- Sample and hold
- Correlated-double-sampling
- FPN suppression

[E. Fossum, 1997]
Architectures for Sensory-Processing

Progressive Distributed Processing

Bio-Inspired Vision Architectures

Shifting the Analog-to-Digital Border
<table>
<thead>
<tr>
<th>Sensors + Analog Signal Conditioning</th>
<th>A/D Conversion</th>
<th>Digital</th>
</tr>
</thead>
<tbody>
<tr>
<td>☺ Large spatial resolution limited only by pixel SNR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☺ Large circuit operation predictability and robustness analog only at the ADCs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☺ Small spurious signal interactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☺ Large flexibility and programmability: mostly digital circuits and codes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☹ Small computational power and efficiency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☹ Large memory requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>☹ Data transfer bottlenecks</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Conceptual Architectural Choices**
**Conceptual Architectural Choices**

**Linear Array of Digital Processors**

- **Smaller spatial resolution**: trade-off with processing
- **Large circuit operation predictability and robustness**: analog only at the ADCs
- **Small spurious signal interactions**
- **Large flexibility and programmability**: mostly digital circuits and codes
- **Larger computational power and efficiency**
- **Smaller memory requirements**
- **Data transfer bottlenecks**

© ANAFOCUS 2008
Conceptual Architectural Choices

A True LAP VSoC for Machine Vision

- Pinned photodiode pixel
- High-speed rolling and global electronic shutter with programmable exposure time.
- Programmable controls: gain, offset, frame rate and frame size (RoI)
- 100MHz clock frequency
- Per-column readout path permitting up to 500fps readout speeds
- Multiple I/Os and high-speed communications
Linear Array of Analog-Digital Processors

- **Sensors**
- **Analog Signal Conditioning**
- **Distributed Analog**
- **A/D Conversion**
- **Distributed Digital: LAP**
- **Digital**

- Smaller spatial resolution: trade-off with processing
- Smaller circuit operation predictability and robustness: analog also for processing
- Larger spurious signal interactions
- Smaller flexibility and programmability: both analog and digital circuits
- Larger computational power and efficiency
- Smaller memory requirements
- Smaller transfer bottlenecks
How Does The Retina Work?

- Sensors and processors are merged
- Processing and sensing are simultaneous
- Significant data compression is achieved
How Does Retina Work?

[Roska & Werblin 2001]
<table>
<thead>
<tr>
<th>Task n°</th>
<th>Processing Task</th>
<th>Example</th>
<th>Amount of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>①</td>
<td>Image Acquiring</td>
<td>Light sensing</td>
<td>b×N^2</td>
</tr>
<tr>
<td>②</td>
<td>Noise reduction</td>
<td>Lowpass, median filtering...</td>
<td>b×N^2</td>
</tr>
<tr>
<td>③</td>
<td>Segmentation</td>
<td>Edge detection</td>
<td>b×N^2</td>
</tr>
<tr>
<td>④</td>
<td>Feature extraction</td>
<td>shape detection</td>
<td>N^2</td>
</tr>
<tr>
<td>⑤</td>
<td>Classification</td>
<td>identification</td>
<td>1</td>
</tr>
</tbody>
</table>


Digital

Digital
Using the Bio-Inspiration Concept

Data Flow

CMOS / CCD sensor

Image Acquisition

High-performance A/D Converters

Image Coding

Advanced DSP & High Density Memory

Image Processing and Storage

Focal-Plane Sensor-Processor

Image Acquisition & Pre-processing

Low-Profile A/D Converters

Pre-processed Image Conversion

Simple DSP & Low Memory

Image Post-Processing and Storage

© ANAFOCUS 2008
Topographic Sensor-Processors

Concept of Concurrent Sensory-Processing

Multi-Functional Pixels

Rationale for Analog Pre-Processing
Concept of Concurrent Sensor-Processing

state/output

input

A

B

Z
Multiple Signal Representations
Basic Spatial Interactions

CNN Single Layer Interactions

FeedForward Connections

Processing Core

FeedBack Connections

\[ y_{cc} = F(x_{cc}) \]

\[ \sum \]

\[ G(\bullet) \]

\[ \int \]

\[ \frac{1}{t} \]
Spatio-Temporal Interactions

CNN Multi-Layer Interactions

Inputs

Layer 1

\( b_1 u_1 + z_1 \)

\( b_2 u_2 + z_2 \)

Inter-Layer Coupling

Layer 2

Feedback

\( \tau_1 \)

\( A_1 y_1 \)

\( a_{12} y_1 \)

\( a_{12} y_1 \)

Layer 3

\( \tau_2 \)

\( A_{22} y_2 \)

\( \tau_3 \ll \tau_1, \tau_2 \)

\( \tau_3 \ll \tau_1, \tau_2 \)

\( w_1 y_1 + w_2 y_2 \)

Outputs

Layer 1

\[ f(\tau_1) \]

\[ g() \]

\[ x_{1,y} \]

\[ y_{1,y} \]

Layer 2

\[ f(\tau_2) \]

\[ g() \]

\[ x_{2,y} \]

\[ y_{2,y} \]

Layer 3

\[ f(\tau_3) \]

\[ g() \]

\[ x_{3,y} \]

\[ y_{3,y} \]
Programming Reaction-Diffusion Equations

Multi-Layer Interactions

\[
\frac{d}{dt} \phi_i(x, y, t) - c_i \nabla^2 \phi_i(x, y, t) = \alpha_i \phi_i(x, y, t) + \beta_i \phi_i(x, y, t_0) + \gamma_{ij} \phi_j(x, y, t)
\]

**diffusion**

**reaction**

Wave propagation in active media

[Perona & Malik 1990]

Layer 1 (slow)

Layer 2 (fast)
Wave Generation

Multi-Layer Interactions

Traveling Waves

Layer 1 (slow)  Layer 2 (fast)

\[
A_1 = \begin{bmatrix}
1 & 1 & 1 \\
1 & 1.5 & 1 \\
1 & 1 & 1 \\
\end{bmatrix} \quad A_2 = \begin{bmatrix}
0.6 & 0.8 & 0.6 \\
0.8 & -0.6 & 0.8 \\
0.6 & 0.8 & 0.6 \\
\end{bmatrix}
\]

\[a_{21} = 4.5 \quad a_{12} = -4.2\]

\[h_1 = 0 \quad h_2 = 5\]

\[z_1 = 0.6 \quad z_2 = -2.7\]

\[\tau_1 : \tau_2 = 16 : 1\]

Spiral Wave

\[
A_1 = \begin{bmatrix}
0.6 & 0.8 & 0.6 \\
0.8 & -0.3 & 0.8 \\
0.6 & 0.8 & 0.6 \\
\end{bmatrix} \quad A_2 = \begin{bmatrix}
0.6 & 0.8 & 0.6 \\
0.8 & -0.9 & 0.8 \\
0.6 & 0.8 & 0.6 \\
\end{bmatrix}
\]

\[a_{21} = 6 \quad a_{12} = -3\]

\[h_1 = 0 \quad h_2 = 3.9\]

\[z_1 = -4.5 \quad z_2 = -3.6\]

\[\tau_1 : \tau_2 = 8 : 1\]
**Multi-Layer Interactions**

**Long Distance Inhibition in the IPL**

Layer 1 (slow)  
Layer 2 (fast)

\[ \mathbf{A}_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1.5 & 1 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 0.6 & 0.8 & 0.6 \\ 0.6 & 0.8 & 0.6 \end{bmatrix} \]

\[ a_{21} = 3.3, \quad a_{12} = -5.4 \]

\[ b_1 = 0, \quad b_2 = 3 \]

\[ z_1 = 0.6, \quad z_2 = -2.7 \]

\[ \tau_1 : \tau_2 = 16 : 1 \]

**Spatial-Temporal Detection of Edges in the OPL**

Only fast layer

\[ \mathbf{A}_3 = \begin{bmatrix} 0.5 & 0.8 & 0.5 \\ 0.8 & -3.6 & 0.8 \\ 0.5 & 0.8 & 0.5 \end{bmatrix}, \quad \mathbf{A}_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \]

\[ a_{21} = 4.5, \quad a_{12} = -5.7 \]

\[ b_1 = 0.0, \quad b_2 = 3.0 \]

\[ z_1 = -9, \quad z_2 = 0 \]

\[ \tau_1 : \tau_2 = 6 : 1 \]
Conceptual Architectural Choices

Topographic ADCs + Digital Processing

- Smaller spatial resolution: trade-off with processing
- Larger system operation predictability and robustness: analog only for ADCs
- Smaller spurious signal interactions
- Larger flexibility and programmability: only digital processing
- Functionality and efficiency compromised by ADC accuracy
- Large in-pixel memory requirements
Conceptual Architectural Choices

Topographic Mixed-Signal Processing

- Low spatial resolution: trade-off with processing
- Involved circuit operation predictability and robustness: analog also for processing
- Larger spurious signal interactions
- Involved flexibility and programability: both analog and digital circuits
- Large computational power and efficiency
- Small memory requirements
- Small transfer bottlenecks
**Rationale for CMOS**

The Functional Power of MOST

- **Ohmic S1-Region:**
  \[ I_D \approx k \frac{W}{L} V_{DS} (V_{G} - V_{T0} - nV_S) \]
  - \( V_{DS} \) must be kept constant
  - Early Voltage is low \( \propto L \)

- **Saturation S1-Region:**
  \[ I_D \approx \frac{k}{2n} \frac{W}{L} (V_{G} - V_{T0} - nV_S)^2 \]
  - \( V_{DS} \) must be kept small

- **Ohmic Region:**
  \[ R_{\Omega} \propto \frac{L}{kW} \]
  - No offset

- **Cut-off Region:**
  \[ R_{ON} \rightarrow \infty \]
  - Charge Injection

- **Strong-Inv. Region:**
  \[ C \propto C_{OX} LW \]
  - Nonlinear

- **Long-Term Memory with Floating Gate**

\[ R_{\square} \sim 10 \text{K} \Omega \]
Adding random noise ($\sigma=4$LSB) to: 1) Input Image; 2) Coefficients of the convolution kernel at every position; 3) Output Image

[Original from Gustavo Liñán]
Using Quality Assessment Methods

The Accuracy Requirements


[Original from Gustavo Liñán]
Multi-Functional Pixel Example

Column

To neighbours

Multipliers

Register

Input Block

From neighbours

Data Bus

Global Control Bus

Local Control Bus

Fast Averager

Binary Operator

Local Masks

Address Event

AE decoder

© ANAFOCUS 2008
The Resolution Compromise

Size of Sensory Pixel

Size of Sensory-Processing Pixel

How Important is Resolution?
How Important is Resolution?

- **512 x 512**
  - Vision is possible with 25 x 25 "pixels" (within limited field of view)
  - Text can be read at 200 words/min (300 words/min with normal vision)
  - Students can navigate in complex environments (maze) with confidence

- **128 x 128**

- **64 x 64**

- **32 x 32**

Resolution can be increased through proper design

- VGA and up to 1.3Mpixels resolution for Surveillance
- > 1Mpixels for Machine Vision and Automotive
The Eye-RIS system

Architecture

The System in Operation
Multi-Functional Pixel Concept

- **F** – Input data flow @ bits per second (bps)
- **f** – Data flow after early processing

Digitization of the pre-processed image

Image post-processing + memory

Actuators

Control signals

© ANAFOCUS 2008
The Eye-RIS™ System

Configurable system which provides an efficient solution for low-to-medium resolution and high frame rates applications

Ultra-high frame rate: up to 10,000fps@QCIF image resolution

Ultra-low power consumption: < 10mW@30fps@QCIF image resolution

© ANAFOCUS 2008
The Q-Eye Chip

- UMC 0.18µm CMOS 1P5M (1.8V/3.3V) - mixed-signal.
- High-performance Smart Image Sensor
  - 176 x 144 grey-scale pixels with 29.1µm pitch
  - High-speed non-rolling electronic shutter. Programmable exposure time (controlling step-down to 20ns)
  - Approximated sensitivity of 3V/lux·sec at 550nm
  - 4 +1 (two banks) high-retention analog and 4 binary memories
  - Analog multiplexer for image shifting and analog MAC unit
  - Programmable, 3 x 3 neighbourhood pattern matching with 1/0/d.n.c. pattern definition (fast morphological functions)
- On-chip bank of high-speed ADCs and DACs.
- Multiple I/Os and high-speed communication ports
Main Features

Q-Eye focal-plane processor
- 176 x 144 spatial resolution
- Monochrome image sensor with 3.2V/(lux·sec) sensitivity
- Maximum frame rate (sensing + processing) of over 10,000fps
- Advanced pixel architecture combining image acquisition, image processing and storage:
  - Multi-mode image sensing, analogue & binary memories, analogue multiplexor for image shifting, analogue MAC unit, programmable LUT, resistive grid for controllable smoothing...

Digital control & post-processing
- ALTERA NIOS-II 32-bit RISC microprocessor
- 1.17 DMIPS/MHz performance at 70MHz operation frequency
- 32Mb SDRAM for program and image/data storage

Communications
- SPI port, UART, 2xPWM ports and GPIOs, USB 2.0, and GigE
- JTAG Controller
- 1.5W typical power consumption

Application development kit
- Application Development Kit including: Project builder, C compiler, assembler, linker, and source-code debugger.
- Image-processing library including basic routines such as: point-to-point operations, spatial filtering operations, morphological operations and statistical operations
The First True Bio-inspired VSoC: Eye-RIS™ v2.1

© ANAFOCUS 2008
Main Features

Q-Eye focal-plane processor
- 176 x 144 spatial resolution
- Monochrome image sensor with 3.2V/(lux·sec) sensitivity
- Maximum frame rate (sensing + processing) of over 10,000fps
- Advanced pixel architecture combining image acquisition, image processing and storage:
  - Multi-mode image sensing, analogue & binary memories, analogue multiplexor for image shifting, analogue MAC unit, programmable LUT, resistive grid for controllable smoothing...

Digital control & post-processing
- ALTERA NIOS-II 32-bit RISC microprocessor
- 1.17 DMIPS/MHz performance at 100MHz operation frequency

Communications
- SPI port, UART, 2xPWM ports and GPIOs, USB 2.0 interface, and JTAG Controller
- <700mW typical power consumption

Application development kit
- Application Development Kit including: Project builder, C compiler, assembler, linker, and source-code debugger.
- Image-processing library including basic routines such as: point-to-point operations, spatial filtering operations, morphological operations and statistical operations
Benchmark: Each of the processors calculated 20 convolutions, 2 diffusions, 3 means, and 40 morphology and 10 global ORs. Note: Only the DSP and the pipe-line multi-core (FPGA) architectures support trading between resolution and frame-rate.
## Processor Comparison

<table>
<thead>
<tr>
<th>Processor</th>
<th>Type</th>
<th>Tech. [nm]</th>
<th>Area [mm]</th>
<th>Data [Mb]</th>
<th>Pow [W]</th>
<th>Clock [MHz]</th>
<th>GOPs</th>
<th>mW</th>
<th>GOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xetal-1</td>
<td>SIMD LPA</td>
<td>180</td>
<td>25</td>
<td>0,05</td>
<td>1,6</td>
<td>24</td>
<td>5</td>
<td>320</td>
<td></td>
</tr>
<tr>
<td>IMAP</td>
<td>SIMD LPA</td>
<td>180</td>
<td>121</td>
<td>2</td>
<td>2</td>
<td>100</td>
<td>51,2</td>
<td>39,06</td>
<td></td>
</tr>
<tr>
<td>Xetal-2</td>
<td>SIMD LPA</td>
<td>90</td>
<td>74</td>
<td>10</td>
<td>0,64</td>
<td>84</td>
<td>107</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Q_Eye</td>
<td>SIMD SPA</td>
<td>180</td>
<td>37,5</td>
<td>Int.</td>
<td>0,1</td>
<td>50</td>
<td>250</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>EyeRIS (NIOS)</td>
<td>SIMD SPA</td>
<td>180</td>
<td>92</td>
<td>0,5</td>
<td>1</td>
<td>50</td>
<td>250</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>CVSoC</td>
<td>SIMD LPA</td>
<td>180</td>
<td>165</td>
<td>0,8</td>
<td>90</td>
<td>92,16</td>
<td>8,681</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Processor Comparison

<table>
<thead>
<tr>
<th></th>
<th>DSP (DaVinci*)</th>
<th>Pipe-line (FP GA**)</th>
<th>Course-grain (Xenon)</th>
<th>Fine-grain (SCAMP/Q-Eye)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silicon technology (micron)</td>
<td>0.065</td>
<td>0.065</td>
<td>0.18</td>
<td>0.35/0.18</td>
</tr>
<tr>
<td>Silicon area mm²</td>
<td></td>
<td></td>
<td>25</td>
<td>50/80</td>
</tr>
<tr>
<td>Power consumption</td>
<td>1.25 W</td>
<td>2.9W</td>
<td>0.08 W</td>
<td>0.20 W</td>
</tr>
<tr>
<td>Arithmetic proc. clock speed</td>
<td>600 MHz</td>
<td>260 MHz</td>
<td>100 MHz</td>
<td>1.2-2.5 MHz</td>
</tr>
<tr>
<td>Number of arithmetic proc.</td>
<td>8</td>
<td>120</td>
<td>256</td>
<td>16384</td>
</tr>
<tr>
<td>Efficiency of arithmetic calc.</td>
<td>75% *</td>
<td>100%</td>
<td>80% **</td>
<td>50% *</td>
</tr>
<tr>
<td>Arithmetic computational speed</td>
<td>3.6 GMAC</td>
<td>30 GMAC</td>
<td>20 GMAC</td>
<td>-20 GOPS***</td>
</tr>
<tr>
<td>3x3 convolution time</td>
<td>42.3 μs****</td>
<td>4.9 μs</td>
<td>12.1 μs</td>
<td>22 μs***</td>
</tr>
<tr>
<td>Arithmetic speed-up</td>
<td>1</td>
<td>8.6</td>
<td>3.5</td>
<td>1.3</td>
</tr>
<tr>
<td>Morph. proc. clock speed</td>
<td>600 MHz</td>
<td>88 MHz</td>
<td>100 MHz</td>
<td>1.2-5 MHz</td>
</tr>
<tr>
<td>Number of morphologic proc.</td>
<td>64</td>
<td>804</td>
<td>2048</td>
<td>147458</td>
</tr>
<tr>
<td>Morphologic processor kernel type</td>
<td>2 x 32 bit</td>
<td>66 x 9 bit</td>
<td>256 x 6 bit</td>
<td>16384 x 6 bit</td>
</tr>
<tr>
<td>Efficiency of morphological calc.</td>
<td>29% *</td>
<td>100%</td>
<td>90% **</td>
<td>100%</td>
</tr>
<tr>
<td>Morphologic computational power</td>
<td>10 GOPS</td>
<td>71 GOPS</td>
<td>184 GOPS</td>
<td>737 GCOPS</td>
</tr>
<tr>
<td>3x3 morphologic operation time</td>
<td>13.8 μs****</td>
<td>2.06 μs</td>
<td>1.1 μs</td>
<td>0.2 μs</td>
</tr>
<tr>
<td>Morphologic speed-up</td>
<td>1</td>
<td>6.6</td>
<td>12.4</td>
<td>68.0</td>
</tr>
</tbody>
</table>

* Texas Instrument DaVinci video processor (TMS320DM6443)
++ Xilinx Spartan 3A DSP FPGA (XC3SD3400A)
* due to data access
** due to pipe-line operations
*** no multiplication, scaling with a few discrete values
**** these data intensive operators slow down significantly when the image does not fit to the internal memory (typically above 128x128 for a DaVinci, which has 64kByte internal memory)

[Original from Csaba Reckeczky]
3-D Parallel Processing Multi-Layer Architectures

Sensor layer (320x240)

Processor array
(160x120)

Mixed signal layer

Control bus

32 bit data bus
(gray value)

Pixel parallel

Bonding wire

Digital proc. layer

Processor array
256 cores

Digital memory layer

memory array
64kB

© ANAFOCUS 2008
Topographic, layered sensing-processing

ADC – Proc No1

Sensor

Proc

Memory

3D CNN CUBE

AFRL BB Process

Bump bonding:

3D stacking:

MITLL Low-Power
FDSOI CMOS Process
Visual Prosthesis

Retina Coder + Electrode Driving + Electrode Array

[K. Wise, D. Kipke, U. Michigan]

Sub-Retinal Implant

[Chow et al. 2001]

Epi-Retinal Implant

[Humayun et al. 1999]
Discussion

- VSoCs are still like science fiction toys for industry
- Even smart CMOS sensors are not common yet
  - A real challenge for engineering !!

- We can take advantage of nature through understanding
  - Parallel and concurrent sensory-processing
  - Multi-scale representation
  - Signal coding

- Close interaction and collaboration between analog and digital circuitry is needed for efficient VSoC design.