RAMP-IV: A Low-Power and High-Performance 2D/3D Graphics Accelerator for Mobile Multimedia Applications

Ramchan Woo, Sungdae Choi, Ju-Ho Sohn, Seong-Jun Song, Young-Don Bae, and Hoi-Jun Yoo

Semiconductor System Laboratory
Dept. of EECS, KAIST
http://ssl.kaist.ac.kr
Outline

• Introduction
  – 3D Graphics on Hands
  – Technical Barriers
  – RAMP Overview

• RAMP-IV Architecture
  – High Performance @ Low Power
  – Low Cost Implementation

• Demonstration
  – REMY : System Evaluation Board
  – MobileGL : Graphics Library for Mobile Devices

• Conclusions
Evolution of the Cell-Phone

- Mobile Multimedia Center
  - Camera
  - 2D Graphics
  - MP3 Audio
  - MPEG-4 Video
  - Java Gaming

- 3DCG is the next step!

Avatar
Game
Advertisement
**Technical Barriers**

- **Limited** Energy supplied by Battery
  - 200mW for 3DCG: 2~3hrs Playback
- **Limited** Resources
  - 400MHz Host Processor without FPU
  - 400MB/s System Memory
- **Limited** Footprint
  - Not Enough Space for 2D/3D Accelerator + Graphics Memories
- **Low** Cost
- **No** Standard 3D Graphics API for Mobile Devices

Low-Power 2D/3D Accelerator integrated with Rendering Memories by DRAM Process Technology → RAMP
What is RAMP?

• RAMP: RAM + Processor
  – Mobile Multimedia Processor with Application Specific Embedded DRAMs
    • High-Level of Integration
    • Low Power Consumption
    • High Performance through Huge Memory Bandwidth
  – Platform integrated with IP Cores

– Merged-DRAM Fabrication Process
  • Embedded Memory Logic (EML) ➔ Pure DRAM

Diagram:

- Fast Cycle Time
- Wide/Separate Bus
- Partial Activation

RISC | Hardware Accelerators (3DRE, MPEG) | Optimized Embedded DRAMs
RAMP History

Hynix 0.35um EML
First Implementation
512kb eDRAM
40Mpixels/s, Shading Only

Hynix 0.18um EML Process
3DCG + MPEG4 on a PDA-Chip
7.125Mb eDRAM
70Mpixels/s, Shading Only

Hynix 0.16um DRAM Process
29Mb eDRAM, Full 3D Pipeline
Texture Mapping, Special Effects
66Mpixels/s, 264Mtexels/s
RAMP-IV Standouts

- **Highest Performance @ Low Power**
  - Highest Integration Level
  - Full 3D Pipeline with Complete Rendering Memories
    - Texture Mapping, Special Effects
    - Large On-Chip Memory
- **Low-Power Techniques**
  - Datapath
  - Shading / Texture Unit
  - Memory Activation
  - Clock Gating / Management
- **Low-Cost Implementation**
  - 256Mb DRAM Process
RAMP-IV Architecture

**ARM9 Core**
- 4KB I$ (Instruction)
- 4KB D$ (Data)
- 1KB DP SRAM

**MAC**
- 32b input
- 128b output

**Clock Control Unit**

**PLL**

**BEQ**
- CTRL
- 1KB DP SRAM
- 32b output

**SlimShader**
- TSE
- PP0
- PP1
- TE0
- TE1
- AAL

**RE**
- Memory Programmer
- 24b output

**GE**
- 32b input

**2Mb Depth Buffer**

**3Mb Frame Buffer**

**24Mb Texture Memory**

**eDRAM**
Geometry Engine

• General Purpose RISC + MAC
  – 133MHz 32bit ARM9 ISA
  – Programmability : No Hard-wired T&L Engine
    • Main Host Processor for Standalone Configuration
  – Enhanced MAC : 32b x 32b MAC in a single cycle
    • 3D Geometry, MPEG Acceleration
    • 1.04Mvertices/s Transformation (43% Improvement)

• Working with Proprietary MobileGL
  – Hand-Optimized for ARM Datapath with Integer MAC
    • Fixed Point Arithmetic
    • x10 Improvement against Floating-Point Emulation
  – Subset of de-facto Standard OpenGL
  – 70kB Library Size
Rendering Engine: SlimShader

- Main Rendering Pipeline
  - 3D Acceleration
  - 2D through 3D Datapath

- Pipeline Clock Gating
  - Early Depth Test
  - Pixel-Level Gating

- Precision-Controlled Datapath
  - 95% Power and 85% Area Reduction
Rendering Engine: SlimShader

- Frame/Depth Buffer
  - Vertical-Stripe Assignment
  - Independent Access Ports
- Per-Triangle Parameter Setup
  - Horizontal-Order Rasterization
    - Easy Memory Addressing
    - Simple Pipeline Control
    - DRAM Page Issues: Solved by 3D-Optimized DRAM
- Two Pixel Processors
  - Energy Efficient: Address Alignment Logic
  - Render Horizontally-Adjacent Pixels
    - Easy to Gather Texture Addresses
    - Simple Assignment
Rendering Engine: Texture Unit

- Two Texture Units
  - Single-Cycle Per-Pixel Dividers for Perspective-Correction
  - Single-Cycle Bilinear MIPMAP Filter

- Address Alignment Logic
  - Energy-Efficient Texturing Unit
  - Attached to 4 Texture Memories
  - Exploiting Locality without using Complex Texture Cache
  - Using Pipeline Latches to store Recently-Used Texels
Benefit of Address Alignment Logic

Required Time

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Required Power

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>4</td>
<td>2.3</td>
<td></td>
</tr>
</tbody>
</table>

Required Energy

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td></td>
<td>2.53</td>
<td></td>
</tr>
</tbody>
</table>

68% Reduction

GE
RE
PPO
eDRAM

Ramchan Woo 13
• Post Processing Unit
  – **Pixel-Programmability without Cycle Penalty**
  – Special Rendering Effects
    • Antialiasing, Fog, Motion Blur
  – Working in Parallel with SlimShader
    • 2-pixel-wide SIMD-Parallel Datapath
    • 160-bit Memory Crossbar
Embedded DRAM

- Memory Configuration

Depth Buffer A0 512kb
Depth Buffer A1 512kb
Depth Buffer B0 512kb
Depth Buffer B1 512kb

Frame Buffer A0 768kb
Frame Buffer A1 768kb
Frame Buffer B0 768kb
Frame Buffer B1 768kb

Texture Memory 0 6Mb
Texture Memory 1 6Mb
Texture Memory 2 6Mb
Texture Memory 3 6Mb

Selective Activation
- Optimized Access
- 75% Power Reduction

Huge Bandwidth
- 416bit
- 1.6GB @ 33MHz

GE  RE
PPO  eDRAM
Embedded DRAM

- Optimized for 3D Graphics
  - Fast Random Row Cycle
    - Matched with Maximum Logic Cycle
    - No Need to Worry about Page-Access
  - Single-Cycle Read-Modify-Write
    - Ideal for Depth-Test, Alpha-Blending
    - Simple Memory Interface

Zero Latency!

\[ tRC = 20\text{ns} \ (50\text{MHz}) \]

- Read Bus
  - Address
    - 10\text{ns}
  - Compare Logic
    - Modify
      - 5\text{ns}
  - Write Bus
    - Write
      - 5\text{ns}
Programmable Power Optimizer

- Run-time Clock Control

- Fully Controlled by the Software
- Adjusting the Frame-Rate during Run-time
- Zero-latency Frequency Change
Low-Cost Implementation

- DRAM-based SoC Implementation
  - Low Cost: 256Mb DRAM
  - "Embedded Processor" into a DRAM-die
  - Large On-Chip Memory
  - Inherently Low-Leakage Current
  - Adequate for Mobile Multimedia Processors
RAMP-IV Characteristics

- 0.16um Hynix DRAM with 1-W 3-Al
- Power Supply
  - 2.0V (eDRAM : Core Transistors)
  - 2.5V (Logic : Periphery Transistors)
  - 3.3V (I/O)
- Power Consumption
  - <210mW (Textured 3D @ FAST)
  - <145mW (Non-Textured 3D @ FAST)
  - < 85mW (MPEG-4 @ FAST)
- Die Size
  - 121mm² (11mm x 11mm)
- Rendering Performance
  - 1Mvertices/s with MobileGL
  - 66Mpixels/s, 264Mtexels/s
  - Perspective-Correct Bilinear MIPMAP Texturing
  - Special Rendering Effects
  - 29Mb Rendering Memories
**Target Configuration**

**Integration with existing A.P.**
- → Highend Cell-Phone

**Replacement of existing A.P.**
- → Low-Cost Cell-Phone

**Attached to main A.P.**
- → Highend PDA
- → Highend Game Terminal

**Standalone Processor**
- → Low-Cost Game Terminal
• System Configuration
  – Standalone Processor

3D Applications

MobileGL

32MByte System Memory
(Samsung Ut-SRAM) or
(Hynix 1T-SRAM)

RAMP-IV

System Interface
(Altera Flex10k)

USB Interface

LCD (256 x 256)

Courtesy of ETRI
REMY: Demonstration Video

• REMY-I
REM My : Demonstration Video

- REMY-II
Conclusions

• RAMP-IV
  – 2D/3D Graphics Accelerator for Mobile Multimedia Applications
  – Highest Level of Integration: GE + RE + DRAM + PPO
  – Highest Performance: 66Mpixels/s, 264Mtexels/s
  – Low Power Consumption: Less than 210mW
  – Low Cost 256Mb DRAM Technology

• REMY
  – Demonstration System
  – MobileGL: Fixed-Point Graphics Library

• 3DCG is Ready for Cell-phones and PDAs!
Related Materials


• Contacts : ramchan@ssl.kaist.ac.kr ; http://ssl.kaist.ac.kr