A Low Power 3D Rendering Engine with Two Texture Units and 29Mb Embedded DRAM for 3G Multimedia Terminals

Ramchan Woo*, Sungdae Choi, Ju-Ho Sohn, Seong-Jun Song, and Hoi-Jun Yoo

Semiconductor System Laboratory
Department of EECS
Korea Advanced Institute of Science and Technology
Outline

• Introduction
• System Architecture
  – Low-Power 3D Rendering Engine
  – Pipeline Structure
  – Texture Unit
• Embedded DRAM
• Implementation Results
• Conclusion
3G Multimedia Terminals

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

- 3DCG is the next step!

Avatar  Game  Advertisement
Rendering Engine Overview

Main Rendering Pipeline

- Triangle Setup Engine
- Texture Unit #0
  - Pixel Processor #0
- Texture Unit #1
  - Pixel Processor #1
- AAL
- Post Processing Unit
  - Texture Memory 0 (6Mb DRAM)
    - TM 1
    - TM 2
    - TM 3
  - Frame Buffer 0 (768Kb DRAM)
    - FB 1
    - FB 2
    - FB 3
  - Depth Buffer 0 (512Kb DRAM)
    - DB 1
    - DB 2
    - DB 3

Vertex Data

Rendered Pixel
Pipeline Architecture with Clock Gating: Low Power (1)

Unnecessary Operation

HOLD #1

HOLD #2

Depth Buffer (eDRAM)

Color/Coordinate Interpolation Unit

Compare

Rendering Clock

Gating Control

ESSCIRC 2003

Ramchan Woo
Triangle Setup Engine

: High Performance

- **Operation**
  - Sorting Vertices
  - Setting-up Parameters
    \[ \Delta(X, Z, R, G, B, U, V, 1/W) / \Delta Y \]

- **Hardware Resources**
  - 3 x 9-way SIMD SUBs
  - 3 x 8-way SIMD DIVs
  - MULs, MUXes

- **Performance**
  - 1 Rendering Cycle (50MHz)
  - ~7,000 RISC Cycles by Software

ESSCIRC 2003

Ramchan Woo 6
Precision-Controlled LUT Divider
: Low Power (2) and Small Area

- Optimized Floating-Point Division
  - Preserving Required Precision
  - Saving 95% Power and 85% Area Compared with IEEE754

Ramchan Woo 2003
Texture Unit: Per-Pixel Division
: Low Power (3) and Small Area

- **Leading Zeros**
- **Meaningless**
- **to 8-bit LUT**
- **Approximation Error**

<table>
<thead>
<tr>
<th>W</th>
<th>Leading Zeros</th>
<th>8-bit Data</th>
<th>LSB</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>U, V</th>
<th>Leading Zeros</th>
<th>8-bit Data</th>
<th>LSB</th>
<th>Zero Padding</th>
</tr>
</thead>
</table>

**Operation**
- \( U = \frac{u}{w}, \) \( V = \frac{v}{w}, \) (where, \( u, v, w = 16\text{bits each} \))
- 16bits / 16bits Division is a Large Overhead
- Use 16bits / 8bits Instead (\( w \geq u, v \))

**Precision Reduction**
- \( 0.78\%_{\text{MAX}} \) Calculation Error
- 95% Area Reduction

ESSCIRC 2003

Ramchan Woo
Texture Unit: AAL Operation

- Low Power (4)

* Bilinear MIPMAP Filtering
  - Requires 4 texels / 1 pixel $\rightarrow$ 8 texels / cycle
* AAL Gathers and Reduces Texel Requests
  - 8 $\rightarrow$ 2.5

ESSCIRC 2003

Ramchan Woo
Address Alignment Logic

: Low Power (5)

Spatial Aligner

Temporal Aligner

• Exploiting Locality
  – Two PPs Render Adjacent Pixels
  – TM Requests : $8 \rightarrow 5 \rightarrow 2.5$

ESSCIRC 2003

Ramchan Woo
Embedded DRAM
: Low-Power (6) and High Performance

• 12 3D-Optimized Memories
  – Selective Activation for Low-Power Consumption

<table>
<thead>
<tr>
<th></th>
<th>Frame Buffer</th>
<th>Depth Buffer</th>
<th>Texture Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{RC}$</td>
<td>20ns (50MHz)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macro Size</td>
<td>768Kbit</td>
<td>512Kbit</td>
<td>6Mbit</td>
</tr>
<tr>
<td>I/O Interface</td>
<td>24bit read</td>
<td>16bit read</td>
<td>24bit I/O</td>
</tr>
<tr>
<td></td>
<td>24bit write</td>
<td>16bit write</td>
<td></td>
</tr>
<tr>
<td>Commands</td>
<td>Read-Modify-Write</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Read, Write</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Auto Refresh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latency</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

• Huge Memory Bandwidth
  – 416bit-Wide On-Die Memory Bus
  – 2.4GByte/s @ 50MHz, Random Row-Change
Integration into a Mobile Graphics LSI

- 0.16um DRAM
  - 1-W 3-AL
- 3DRE + eDRAM
  - 46mm²
  - 5M ~ 50MHz Scalable
  - 10M ~ 100Mpixels/s
  - 40M ~ 400Mtexels/s
- Mobile Graphics LSI
  - 121mm²
  - Chip : 33MHz
  - 210mW (Full Die)
  - 140mW (3DRE+eDRAM)
REMY-I : First Prototype System
REMY-II : PDA Prototype
Summary

• Low-Power 3D Rendering Engine for 3G Multimedia Terminals
  – (1) High Performance
    • x50 Faster than QVGA Phone Requirements
  – (2) Low Power Technique
    • <140mW
  – (3) Low Cost Implementation
    • 0.16um 256Mb DRAM Process Technology

• <140mW, 46mm$^2$
  – Integrated into the Mobile Graphics LSI
  – Successfully Demonstrated on System Evaluation Boards