A Fixed-point Multimedia Co-processor with 50Mvertices/s Programmable SIMD Vertex Shader for Mobile Applications

Ju-Ho Sohn, Jeong-Ho Woo, Ramchan Woo, and Hoi-Jun Yoo

Semiconductor System Lab.
Dept. of EECS
Korea Advanced Institute of Science and Technology
Outline

• Introduction
• System Architecture
• Low Power Features
  – Fixed-point processing
  – Instruction-wise clock gating
• SIMD Datapath Structure
  – ALU, Multiply, SFU
• Implementation Results
• Summary
3G Mobile Terminals

- **Requirements**
  - Low power (< 200mW)
  - High performance (>1Mvertices/s)
  - Variable Applications (Programmability)
  - Low cost (<5$)

- **Fixed-point multimedia co-processor**
  - Low power consumption in ARM-10 architecture
  - Programmability with SIMD vertex shading

Wireless Multimedia Center

- MP3 Audio
- Java Game
- MPEG Video
- 3D Graphics

MP3 Audio
Java Game
MPEG Video
3D Graphics
Programmability in Graphics Hardware

Traditional Graphics Pipeline

- Transform
- Lighting
- Clipping
- Rasterization
- Texture Mapping

Programmable Graphics Pipeline

- Vertex Shader

User-defined Vertex Processing

Flexible 3D Graphics & Multimedia Functions

Fixed Functions

Simple Graphics

ESSCIRC 2005
Co-processor Architecture

- 128-bit 4-way SIMD ARM-10 Co-processor
Dual Operations

- Two operating states
  - Tightly coupled co-processor (TCC)
    - Normal co-processor
    - General SIMD instructions
    - Handshaking with ARM-10
  - Parallel processor (PP)
    - Independent processor
    - Graphics instructions
    - Parallel operations with ARM-10
Processor State Control

- Enhanced co-processor I/F for PP state
  - VPCTRL with 2kB code memory
  - Share all HW blocks with TCC state
  - Drive CPBusy for synchronization
Programmable Vertex Shader

- Shader instruction extensions in PP state
  - Source swizzling and write masks
    \[ VGR2 = VGR0 + VGR1 \quad VGR2.xyzw = VGR0 + VGR.xyzw \]
      (TCC state) \quad (PP state)
  - More operands
    \[ VGR2 = VGR0 + VGR1 \quad VOR2.xyzw = VIR0 + MEM[1] \]
      (TCC state) \quad (PP state)

- Parallelism for streaming graphics data

\[ \text{External Memory} \quad \text{ARM-10} \quad \text{Co-processor} \]

1st Vertex Fetch \quad 2nd Vertex Fetch

1st Vertex Processing \quad 2nd Vertex Processing

Hide external memory transfer cycles

Shading

ESSCIRC 2005

Ju-Ho Sohn
Fixed-point Processing
: Low Power Features (1)

- **Fixed-point number representation**

  Bit index: $b_7 b_6 b_5 b_4 b_3 b_2 b_1 b_0$, $b_i \in \{0,1\}$

  Value:
  
  $2^3$ $2^2$ $2^1$ $2^0$ $2^{-1}$ $2^{-2}$ $2^{-3}$ $2^{-4}$

  (ex) Q4.4 format

- **Energy efficiency**
  - 40% less energy than floating-point on avg. in 3D graphics
Instruction-wise Clock Gating
: Low Power Features (2)

- SIMD RF and datapath are the most power-consuming parts
  - Activation in instruction-by-instruction basis

```
    Clock Source
    Enable

    ARM-10 -> Co-Proc I/F -> Valid
    Reg. Files -> Fixed-point SIMD Datapath

    Multimedia Co-processor

    Clock-off

    High only @ CP is called
    Prevent signal transitions
```
**SIMD Datapath : ALU**

- Single-cycle integer and fixed-point op.
  - Arithmetic, logic, shuffle, packing, alignment
  - SW floating-point emulation (80Mflops peak)
SIMD Datapath: Multiply

- 4-cycle matrix transformation

Broadcasted Vector Elements

Convert 32b Fixed to 64b Integer

Accum.

MUL

32x16 MUL

32x16 MUL

SHIFT

CSA

CPA

CSA

CPA

SHIFT

Bypass of Low 32b

Bypass of High 32b

Convert 64b Integer to 32b Fixed

Fixed-point Result

ESSCIRC 2005

Ju-Ho Sohn
SIMD Datapath: Special Function Unit

- Calculate RCP (1/x) and RSQ (1/sqrt(x))
  - Fixed-point input → Fixed-point output
  - No additional data type conversion circuits

*For Qm.n Fixed-point Format*

```
<table>
<thead>
<tr>
<th>OP</th>
<th>Pre-scale Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCP</td>
<td># of Leading Zero</td>
</tr>
<tr>
<td>RSQ</td>
<td>n/2 + # of Leading Zero</td>
</tr>
</tbody>
</table>
```

- Scale up fixed-point dividend to 64b space
Integration into a Mobile Graphics Processor

- **0.18μm CMOS**
  - 1P6M

- **200MHz**

- **10.2 mm²**
  - Full chip (36 mm²)

- **75.4 mW**
  - Full chip (155mW)

- **3D performance**
  - 50Mvertices/s
    (Parallel projection)
  - 3.6Mvertices/s
    (Full 3-D geometry)
System Evaluation Board
Conclusion

• Low power multimedia co-processor for 3G wireless terminals
  – Programmable SIMD vertex shader
    • ARM-10 co-processor architecture
    • Flexibility for various multimedia applications
  – Low power features
    • Energy-efficient fixed-point SIMD datapath
    • Instruction-wise clock gating
  – High performance
    • 50Mvertices/s for parallel projection

• <75.4mW, 10.2 mm²
  – Integrated into the mobile graphics processor
  – Successfully demonstrated on evaluation board