Frame Buffer Design for Image Sensor Array

ECE404: Semester Project

Raj Parihar
parihar@ece.rochester.edu
Motivation

- Image sensors generate large bit-streams
  - High bandwidth is needed to transmit the data
  - Centralized server/DSP is always engaged
  - Server consumes lot of power in processing the raw data

- An alternative is to pre-process the data at sensors
  - Popularly known as pixel-level processing
    - For example, image sensor can provide the data which is either accumulated or transformed as oppose to bit-stream

- How feasible/ beneficial is to implement such ideas?
Outline

- Image Sensor Overview
- Frame Buffer Organization
- Monolithic Implementation
- Distributed Implementation
- Two-Level of Accumulation
- Simulation and Analysis
- Conclusion and Recommendation
Objectives

- Minimize the overall power consumption of system
  - System = Image Sensor Array + Frame Buffer + Additional Support Circuitry

- Estimate the power consumption for DCT and Wavelet pre-processing

- Choices to implement frame buffer
  - Low power DRAM to save power
  - Additional circuitry to reduce the memory references
*Source: “0.88nw/pixel, 99.6 dB linear dynamic range fully digital Image sensor employing a pixel level Sigma-Delta ADC”, Z. Ignjatovic, Mark F. Bocko
**Sigma Delta Modulator**

*Source: “Principle of Sigma Delta Modulator for A/D converter”, Sangil Park*
Frame Buffer: Raw Accumulation

1024 x 1024 image sensor

MUX Adder Pair

1024 bytes

1 MB Storage
DCT Image Accumulator Architecture

1024 x 1024 image sensor

8:1

DCT/ WAVELET Hardware

Coefficient Memory

64B buffer

128 bytes

1 MB Storage

1024 rows

3/7/2012
Basic Building Blocks

- **Image Accumulator Architecture**
  - 1 MB of data array
  - Adders/ Counters
  - Row multiplexers

- **DCT/ Wavelet based Accumulation**
  - DCT Hardware
  - Additional coefficient buffers
Simulation Environment

- Simulation Tool
  - CACTI 5.3

- Technology
  - 90 nm

- Supply Voltage
  - DRAM Data Array: 0.9 V
  - SRAM Data Array: 1.2 V

- Technology cells
  - ITRS – HP, ITRS – LOP, COMM – DRAM
CACTI 5.3: Timing Model

- **Random Access Time (RAT)**
  - Time taken to get the data from storage cell to output of the sense amplifier – into row buffers

- **Random Cycle Time (RCT)**
  - Time taken to activate any row in the same bank

- **Multibank Interleave Cycle Time (MICT)**
  - Time taken to activate any row in other bank

- **Total Access Time**
  - RAT (Same bank and Same row)
  - RAT + MICT (Different bank)
  - RAT + RCT (Same bank Different row)
Constraints and Requirements

- Overall Power Consumption
  - Order of Image sensor power consumption
  - 1024 x 1024 size image sensor consumes < 1mW
  - Power Budget: < 10 mW

- Storage Requirement
  - 1024 x 1024 pixel requires 1 MB of storage

- Bandwidth Requirement
  - High quality video – 30 Frames/ Second
  - Total Bandwidth = Frames/Sec * Size * OSR * 2
  - Total BW = 30 x 1024 x 1024 x 256 x 2 ~ = 16 GBps
    - If the resolution is 12 bit then we require 24 GBps of BW
High Throughput

There are various ways to achieve high bandwidth

- Single bank with wider data port
- Many banks in parallel
- Pipelining of Operations
- Multiple devices in single Rank
- Larger row or page size

Some of these increase the power consumption too

- The idea is to increase the BW without increasing power consumption tremendously
Single SRAM Array

- Power consumption for single SRAM array of 1 MB for 16 GBps bandwidth

![Graph showing power consumption for SRAM array](image.png)
Single DRAM Array

- Power consumption for single DRAM array (1 MB) for 16 GBps bandwidth

![Power Consumption Graph]

**Best:** 162 mW
Monolithic Implementation

- 1 MB of Frame Buffer is implemented as a single device to provide the 16 Gbps of BW
  - SRAM based implementation
    - POWER: ~ 850 mW
  - DRAM based implementation
    - POWER: ~ 160 mW

Conclusion
- On an average, SRAM array is 5-times more power hungry than equivalent DRAM array
- This is the reason why DRAM is preferred over SRAM
Multiple Bank Organization

- The idea is to find out the optimum array size with data width which
  - Minimizes the power consumption
  - Maximizes the bandwidth

- SRAM based solution
  - SRAM is at least 5 times more power consuming
  - Power is primary concern so SRAM is not viable option

- DRAM based solution
  - Smaller DRAM array are comparable with the SRAM performance in terms of power and BW
  - Find the appropriate DRAM organization
DRAM System: Higher Rank

- Mainly used to provide the wider data width

BW Requirement = 8 GBps; Why?
Two-Rank Implementation

- Idea is to overlap the READ and WRITE
  - Total Capacity per rank: 512 KB
  - Bus width per rank: 1024

64 – KB means
8 DRAM Devices of 64 KB
Each with 128 – bit data wide

![Power Consumption Graph](https://via.placeholder.com/150)

Two - Rank: 1 MB Storage for 16 GBps

Array Size (KB)

<table>
<thead>
<tr>
<th>Array Size (KB)</th>
<th>Total Power Consumption (mW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>201.3</td>
</tr>
<tr>
<td>8</td>
<td>173.0</td>
</tr>
<tr>
<td>16</td>
<td>127.7</td>
</tr>
<tr>
<td>32</td>
<td>163.6</td>
</tr>
<tr>
<td>64</td>
<td>119.0</td>
</tr>
<tr>
<td>128</td>
<td>155.3</td>
</tr>
<tr>
<td>256</td>
<td>159.6</td>
</tr>
<tr>
<td>512</td>
<td>185.4</td>
</tr>
</tbody>
</table>
Two-Level of Accumulation

- **Second Level**
  - Larger DRAM array which serves the capacity
  - Less often accessed so low dynamic power

- **First Level**
  - Smaller DRAM/ SRAM array
  - Provides the aggressive bandwidth
  - Data is transferred to level-2 after accumulation
First-Level Buffer

Array Size (KB) and Data Width (bits)

Power Consumption (mW)

- 4 KB – 512b (14mW)
- 8 KB – 1024 b (16mW)
Second-Level Frame Buffer

- Total 4 Banks – each with 256 KB size
- Two bank would contribute to dynamic power whereas other two would incur the leakage power

![2nd Level Buffer: DRAM of 256 KB](chart.png)
Two-Level DRAM Buffers

@ 8 MHz

@ 128 MHz

1024 x 1024 image sensor

64, 8-bit Gray counters

64, 16:1 MUX

16:1

@ 128 MHz

512KB DRAM

2048

@ 64 MHz

512KB DRAM

1 mW

1024 rows

1024 x 1024 image sensor

1024 x 1024 image sensor

1 mW

1024 x 1024 image sensor

1 mW
Two-Level Accumulation

- First Level Buffer
  - 4, 4KB – 512 bit wide DRAM Array
  - Total Power Consumption = 14 mW

- Second Level Buffer
  - 4, 256 KB – 2048 bit wide DRAM Array
  - Total Power Consumption = 26 mW

- Total Power Consumption
  - Approximately 40 mW
Comparison

Various Frame Buffers

<table>
<thead>
<tr>
<th>Implementation</th>
<th>Power Consumption (mW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRAM based</td>
<td>847</td>
</tr>
<tr>
<td>DRAM based</td>
<td>162</td>
</tr>
<tr>
<td>MultiRank DRAM</td>
<td>119</td>
</tr>
<tr>
<td>Two-level DRAM</td>
<td>40</td>
</tr>
<tr>
<td>Pixel-Level 1T DRAM</td>
<td>10*</td>
</tr>
</tbody>
</table>

3/7/2012

Image Accumulator Architecture
Raj Parihar
Optimizations & Recommendations

- Use of gray counters in place of adders or normal counters
- Use of gray coding in address
  - Less switching than binary
- Design low power MUX using pass logic
- Row data interleaving
  - Reduces the number of accumulator required
  - Reduces the width of data bus of data array
- Incorporation of bit-wide memory into pixels
Row Data Interleaving

- Column select

Otherwise 8192 wide port is required as oppose to 1024-bit wide

Reduces the data width
Reduces the number of Adders/Counters
Increases the operational frequency

![Diagram showing row data interleaving with column select and 8:1 multiplexer]

7, 6, 5, 4, 3, 2, 1, 0

S0, S1, S2

Otherwise 8192 wide port is required as oppose to 1024-bit wide
Low Power Multiplexers

Roughly 50% power saving compare to CMOS implementation
N Bit-wide Pixel Memory

- Pixel memory cell could be a 3T or 1T based DRAM cell
- Power consumption of data arrays are cut by a factor of $2^n$
  - For example, 2-bit memory would reduce the power by 4 times
- Every pixel would require an individual DAC + 2-bit Memory cell
Pixel-Level Memory

Low Resolution Algorithm

- Proposed algorithm to implement reduced resolution!
- If server needs a reduced resolution image
  - Server doesn’t require to do the adjustment
- The quality is still quite good: Tested using MATLAB
From 90 nm to 65 nm

- Dynamic power consumption @ $F_{\text{MAX}}$
  - Increases by ~1.5x compared to 90 nm

- Maximum operating frequency increases by ~ 1.7x
  - This means the whole system can be slowed down by 1.7 times and still achieve the same performance
  - Power = $K \times (\text{Frequency})^3$

- Net power saving = $(1.7)^3/1.5 = \sim 3x$
Future Scope

- Power estimation and Design of
  - Row multiplexers
  - Low power arithmetic blocks
  - Memory controller design

- Design of DCT and wavelet hardware
  - Power estimation of design in 65 nm technology
Conclusion

- DRAM based Systems
  - Go wider, rather than Faster

- In DRAM systems primary focus is to achieve higher bandwidth rather than low latency

- Wider DRAM structure provide more bandwidth
  - Can operate with slower clock in order to save power, if required
Backup Slides
DRAM Arrays with 4 Banks
### First Level: Power Consumption (4-Bank)

<table>
<thead>
<tr>
<th>Bus Width (b)</th>
<th>4 KB</th>
<th>8 KB</th>
<th>16 KB</th>
</tr>
</thead>
<tbody>
<tr>
<td>128 -</td>
<td>35.8</td>
<td>46.0</td>
<td>54.0</td>
</tr>
<tr>
<td>256 -</td>
<td>31.3</td>
<td>59.6</td>
<td>37.4</td>
</tr>
<tr>
<td>512 -</td>
<td><strong>13.9</strong></td>
<td>15.9</td>
<td>46.3</td>
</tr>
<tr>
<td>1024 -</td>
<td>15.7</td>
<td>15.7</td>
<td>38.6</td>
</tr>
</tbody>
</table>