Energy and Area-Efficient Hardware for HEVC Inverse Transform

Mehul Tikekar, Chao-Tsung Huang, Vivienne Sze, Anantha Chandrakasan

Massachusetts Institute of Technology



Inverse Transform -H.264/AVC vs. HEVC

| | H.264 /AVC | HEVC | Implementation Challenges |
|---------------------------------|---------------|---------------------------|---|
| Transform units (TU) variety | 4x4, 8x8 | 4x4, 8x8, 16x16, 32x32 | Complex pipelining |
| Largest TU size | 8x8 | 32x32 | 4x computation per pixel 16x memory requirements |
| Transform precision | 5-bit | 8-bit | 2x multiplier logic |
| Transform types | IDCT | IDCT, IDST (4x4 only) | |
| Software run time | < 11% | 12% - 23% | |

Inverse Transform



Inverse Transform in Hardware





Hardware metrics and Contributions

- Energy per pixel
 - Depends on statistics of input data
 - Propose: data-gating in 1-D transform
- Area
 - Depends on throughput, transpose memory size
 - Propose: SRAM-based transpose memory
- Throughput
 - Target: 4K Ultra-HD 30fps = 400 Mpixel/cycle
 - 2 pixel/cycle at 200 MHz
 - Propose: zero-coefficient column skipping, register cache for transpose memory

1-D Transform Logic – Partial Butterfly Structure



Spurious Switching Activity for IDCT-4



Energy Savings by Data-gating

| IDCT size | Energy Savings | |
|-----------|----------------|--|
| 4-pt | 37% | |
| 8-pt | 31% | |
| 16-pt | 9% | |
| 32-pt | -12% | |

3% - 26% savings over all quantization parameters and encoding configs in JCT-VC common test condition







































































- 16kb memory:
 - With register array: 125 kgates (a complete H.264/AVC decoder area)
 - SRAM-based design for low area cost

| | 1-port SRAM | Register array | |
|---------------------|---------------------|------------------------|--|
| Transistors per bit | 6 | 30 | |
| Access flexibility | Low (address based) | Arbitrary access | |
| Throughput | 1 entry per cycle | Entire array per cycle | |
| Read latency | 1 cycle | 0 cycle | |

Interleaved Addressing for Transpose Memory

- 4 SRAM banks
- Each SRAM entry stores 1 pixel



8x8 Transform unit



Pipeline stall due to SRAM



Register cache to remove stall



Zero-column skipping – Motivation



Zero-column skipping



- Save 39% cycle count (input data dependent)
- Save 27% energy per pixel (reduced clocking, SRAM writes)

Implementation results

| Designs | Logic area (kgate) | Energy per pixel (pJ/pixel) | Throughput (pixel/cycle) |
|------------------|-----------------------|--------------------------------|-----------------------------|
| Base design | 118 | 18 – 32 | 2.0 |
| Zero-column skip | 122 | 13 – 30 | 2.3 – 3.5 |
| Data-gating | 123 | 18 – 25 | 2.0 |
| Complete design | 126 | 12 – 22 | 2.3 – 3.5 |

- 43% energy savings
- 50% throughput improvement
- 7% area increase
- Energy computed from post-layout simulation
- Energy and throughput measured under JCT-VC common test conditions1

Data-dependent Energy/pixel



Summary

- HEVC Inverse Transform requires 8x computation per pixel and 16x memory as H.264/AVC which increases energy/pixel and area
- This work proposes:
 - Data-gating to reduce energy/pixel by 17%
 - SRAM-based transpose memory to reduce area
 - Register cache for transpose memory to increase throughput
 - Zero-column skip to reduce energy/pixel by 27% and increase throughput by 39%