AN ARCHITECTURE STUDY OF BILATERAL FILTERS (JVET-L0049)

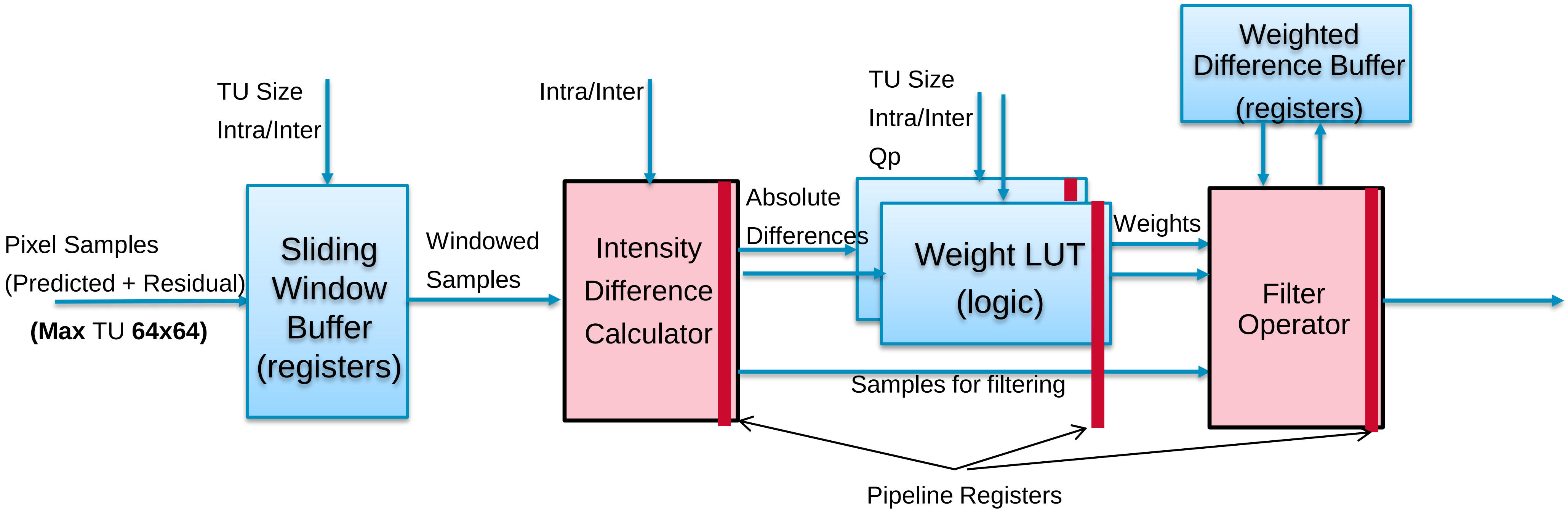
Yi Hu and Minhua Zhou

Broadcom Inc.

Introduction

- Describe a hardware architecture for implementing the in-loop bilateral filter proposed in JVET-K0384 and JVET-K0274
- Estimate the hardware complexity and throughput rate for three configurations
  1. Configuration A: 1 sample/clock cycle
  2. Configuration B: 4 samples/clock cycle
  3. Configuration C: 8 samples/clock cycle
- Hardware complexity (silicon area) is estimated by synthesizing prototype RTL using Design Compiler, measured in equivalent gate count (2-input NAND gate count). The design library is TSMC 28nm standard cell library.
- The analysis is aiming for providing a better understanding of bilateral filters in terms of implementation cost
  - Not for endorsement or criticism of any design on the table

In-loop Bilateral Filter HW Architecture



Pixel sample I/O format

- Pixel samples are fed to the filter block by block and line by line within the block,
  - Configuration A: 1 sample per clock cycle
  - Configuration B: 4 samples per clock cycle
  - Configuration C: 8 samples per clock cycle

Filtering Latency

- Sliding window buffer introduces variable delays depending on block sizes and configuration
- The Filter needs 3 pipeline stages – 3 clock cycle latencies

Sliding Window Buffer

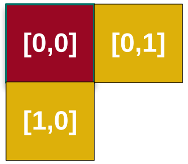
- ❑ **Sliding window buffer provides the sample for filtering and the neighboring samples for calculating the filter coefficients(weights)**
- ❑ **It is determined by the block size and number of samples processed in parallel**

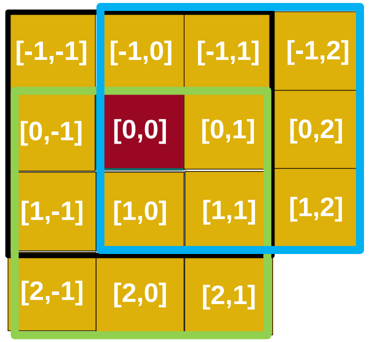
- The neighboring samples at the left and the above of the filtering sample are not directly stored in the buffer
- Number of samples in the buffer is dependent on configuration and block type.
- Maximum intra block size is 64x64
- Inter block needs to buffer more because of 3x3 average of absolute sample difference

Filter Configuration	Configuration A	Configuration B	Configuration C
Max. number of samples for Intra block	65	68	72
Max. number of samples for Inter block	195	201	209
Gate count	16.1K	17.7K	19.3K

- JVET-K0384 and JVET-K0274 require the same sliding window buffer

Intensity Difference Calculator





Two intensity differences, $Diff_{right}$ and $Diff_{below}$, **are calculated for each pixel filtering**

- For intra block, $Diff_{right} = abs(p[0,1] - p[0,0])$, $Diff_{below} = abs(p[1,0] - p[0,0]$
- For inter block, $Diff_{right} = \frac{N}{M} \sum_{i=-1}^{+1} \sum_{j=-1}^{+1} abs(p[i, j+1] - p[i,j])$, $Diff_{below} = \frac{N}{M} \sum_{i=-1}^{+1} \sum_{j=-1}^{+1} abs(p[i+1, j] - p[i,j])$
- For configuration B and C, some intermediate results (the absolute differences) can be shared between filtering samples

Filter Configuration/ Gate Count	Configuration A	Configuration B	Configuration C
JVET-K0384, N=1, M=4	3.9K	9.7K	13.5K
JVET-K0274, N=114, M=1024 (N/M~=1/9)	4.3K	10.9K	15.4K

Weight (Filter Coefficient) LUT

To filter one pixel, four weights(filter coefficients) are required. Two weights, W_{right} and W_{below}, are directly generated by LUT(combinational logic). The other two, W_{left} and W_{above}, are obtained by buffering W_{right} and W_{below} in Weighted Difference Buffer.

In JVET-K0384, the weight is defined as W $= Distance(blk\_size, blk\_type) * Range(Diff, blk\_qp)$. Each LUT has 28.8K non-zero entries and each entry has 18-bit precision. This requires large amount of gates to implemented. Reduce the weight precision can reduce the number of non-zero entries and the required gate count. The table below lists the gate count for 18-bit and 12-bit precision.

In JVET-K0274, the weight is defined as W $= Range(Diff, blk\_qp)$. The multiplication by $Distance(blk\_size, blk\_type)$ is moved into the Filter Operator. $Range(Diff, blk\_qp)$ is approximated by selecting one of four 8-bit LUTs, $Range(Diff, 50)$, $Range(Diff, 46)$, $Range(Diff, 42)$ and $Range(Diff, 38)$ with a shift operation. The weight precision is equivalent to 12-bit.

Weight Precision	Number of non-zero entries	one Weight Gate Count	Configuration A Gate Count (2 Weights)	Configuration B Gate Count (8 Weights)	Configuration C Gate Count (16 Weights)
JVET-K0384: 18-bit	28,795	23.7K	47.3K	189.8K	389.6K
JVET-K0384: 12-bit	23,725	14.8K	29.7K	119.0K	238.0K
JVET-K0274: 12-bit	769	1.8K	3.6K	14.4K	28.8K

Filter Operator and Weighted Difference Buffer

Performing the following operations –

- $WD_{right} = W_{right} * (p[0,1] - p[0,0])$ and $WD_{below} = W_{below} * (p[1,0] - p[0,0])$
- Buffer WD_{right} for use as WD_{left} in pixel p[0,1] filtering, and buffer WD_{below} for use as WD_{above} in pixel p[1,0] filtering
- For JVET-K0384, the filtered pixel sample $p'[0,0] = \left(p[0,0] + WD_{right} + WD_{below} - WD_{left} - WD_{above} + (1 \ll 19)\right) \gg 18$
- For JVET-K0274, the filtered pixel sample $p'[0,0] = Distance(blk\_size, blk\_type, Boundary\_cond) * \left(p[0,0] + WD_{right} + WD_{below} - WD_{left} - WD_{above} + (1 \ll 17)\right) \gg 16$, here $Distance(blk\_size, blk\_type, Boundary\_cond)$ is an 8-bit unsigned value generated by a small LUT, $Boundary\_cond$ indicates the pixel is inside, or at the edge or corner of the block

Filter Module	Configuration A Gate Count	Configuration B Gate Count	Configuration C Gate Count
JVET-K0384: Weighted Difference Buffer	14.8K	15.1K	15.5K
JVET-K0384: Filter Operator	3.2K	12.8K	25.6K
JVET-K0274: Weighted Difference Buffer	10.3K	10.6K	11.0K
JVET-K0274: Filter Operator	3.5K	14.0K	28.0K

Summary – Estimated Gate Count

Proposal	JVET-K0384			JVET-K0274		
Configuration	Configuration A	Configuration B	Configuration C	Configuration A	Configuration B	Configuration C
Sliding Window buffer	16.1K	17.7K	19.3K	16.1K	17.4K	19.3K
Intensity Difference Calculator	3.9K	9.7K	13.5K	4.3K	10.9K	15.4K
Weight LUT	47.3K	189.8K	389.6K	3.6K	14.4K	28.8K
Weighted Difference Buffer	14.8K	15.1K	15.5K	10.3K	10.6K	11.0K
Filter Operator	3.2K	12.8K	25.6K	3.5K	14.0K	28.0K
Total	85.3K	245.1K	463.5K	37.8K	67.3K	102.5K

Notes

- JVET-K0274 has optimized Weight LUT but lower computation accuracy
- JVET-K0384 does not have all the computation specified in bit-exact manner. The Weight LUT precision can be reduced to reduce the gate count.