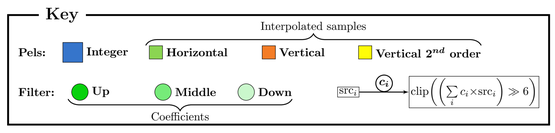
Coding- and Energy-Efficient FME Hardware Design

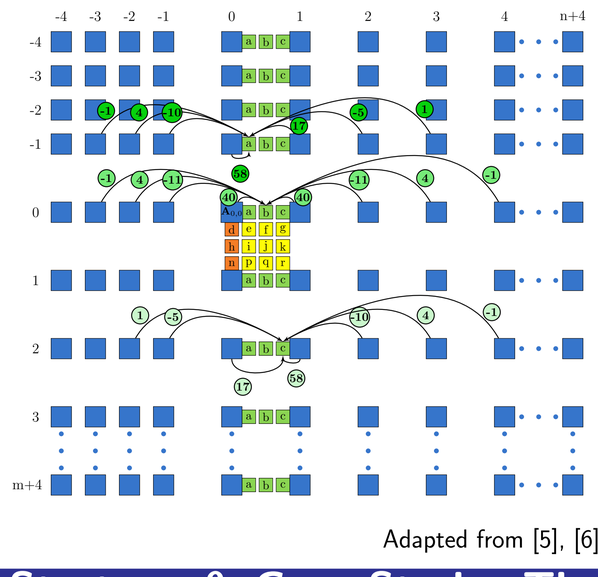
Ismael Seidel[†]*, Vanio Rodrigues Filho[†], Luciano Agostini*, José Luís Güntzel[†]

[†]Embedded Computing Lab. – UFSC, Florianópolis, Brazil, *Video Technology Research Group (ViTech) – UFPel, Pelotas, Brazil

1. Introduction & Definitions

- In 2016, 73% of all Internet traffic was video [1]
- Thus there is a need for ever higher compression
- High Efficiency Video Coding (HEVC) [2] improves ∼50% the coding efficiency over its predecessor
  - But at a huge cost in computational complexity [3]
- Fractional Motion Estimation (FME) is responsible for up to $\mathbf{33.88\%}$ of inter prediction time in HEVC Model [4]
  - $\mathbf{46.9\%}$ of FME time is spent interpolating pixels
  - $\mathbf{53.1\%}$ is spent doing block matching





Adapted from [5], [6].

- Block Matching Algorithm (BMA) [7]:

$$\boldsymbol{B}^{\text{ref}} = \underset{\boldsymbol{B}^{\text{can}} \in \boldsymbol{S}}{\arg\min}\, \boldsymbol{cost}_{m\times n}(\boldsymbol{B}^{\text{ori}}, \boldsymbol{B}^{\text{can}}) \quad (1)$$

- Rate-Distortion Optimization (RDO): Lagrangian rate-distortion cost (j_{cost}) [8] as $\boldsymbol{cost}$:

$$j_{\text{cost}}(\boldsymbol{B}^{\text{ori}}, \boldsymbol{B}^{\text{can}}) = \boldsymbol{distortion}(\boldsymbol{B}^{\text{ori}}, \boldsymbol{B}^{\text{can}}) + \boldsymbol{\lambda} \times \boldsymbol{rate}(\boldsymbol{B}^{\text{ori}}, \boldsymbol{B}^{\text{can}}) \quad (2)$$

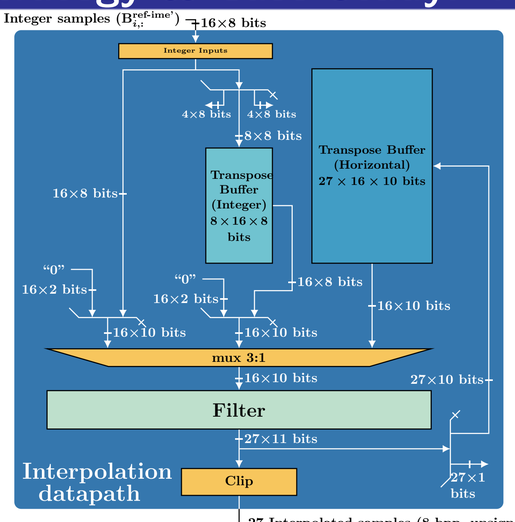
$$\text{sad}_{m\times n}(\boldsymbol{B}^{\text{ori}}, \boldsymbol{B}^{\text{can}}) = \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} |\boldsymbol{d}_{i,j}| \quad (3)$$

2. Proposed Design Strategy & Case Study: The HEVC 8×8 FME

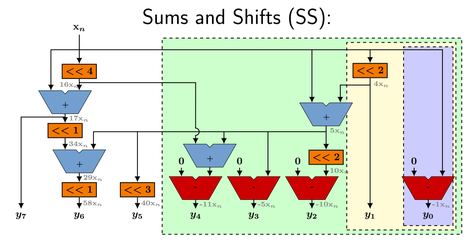
Contributions

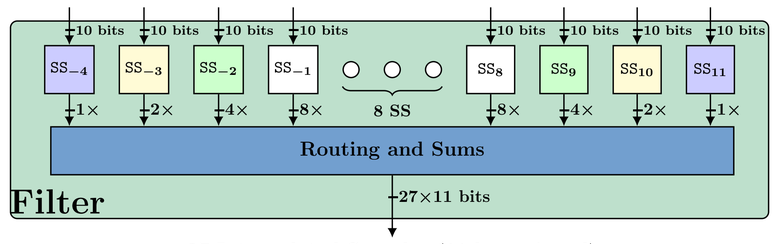
- **Main: Design Strategy!**
  - **Coding-efficient**
  - **Energy-efficient**
- Secondary:
  - Area and Power/Energy Breakdown
- More details in our paper
  - Enjoy open-preview:

- **Memory-Aware**
  - Loads all the required pixels only once
  - Keeps data internally
  - Does transposition internally

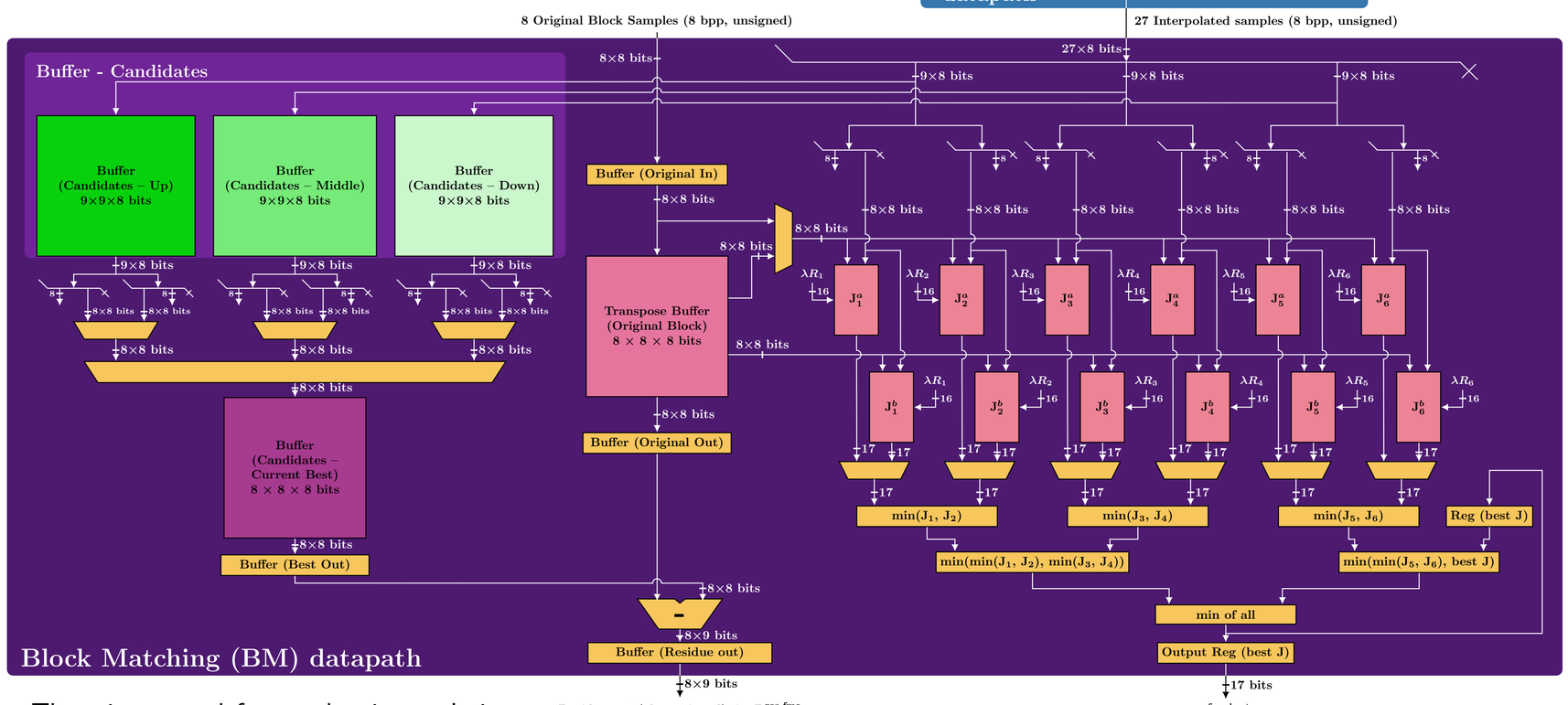


- Using Multiplierless Constant Multiplication (MCM)
- No redundant operations
  - Exploring common sub-expressions





- **Interpolation** and **BMA** integrated in the same hardware
- Using data as soon as possible



- Searching all 48 FME candidates

$$\mathbf{6 + 6 + 3 \times 12}$$

Horizontals+Verticals+Verticals 2^{nd} order

- Leads to a regular dataflow
  - Allows to explore data-level parallelism
- BD-Rate:
  - **-0.37%** (RA)
  - **-0.77%** (LD)

- Using $\boldsymbol{\lambda} \times$ rate (j_{cost})
  - **-1.43%** (RA)
  - **-0.86%** (LD)

- There is no need for another interpolation:

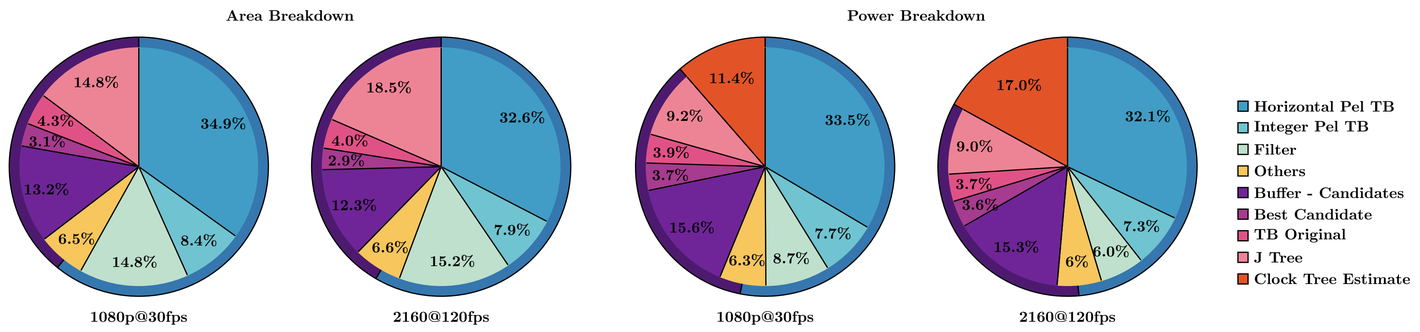
3. Area and Power Evaluation

- The case study architecture was synthesized with Synopsys® Design Compiler (DC®) [9]

Synthesis results for TSMC 45 nm @ 0.9 V standard cell library [10]

Target	1080@30fps	2160@30fps	2160@60fps	2160@120fps*
Period (ns)	20	5	2.5	1.25
Area (μm^2)	93738.3	93826.5	94439.1	100733.4
Dynamic (mW)	3.05	10.78	21.21	44.04
Static (mW)	0.91	0.91	0.92	1.04
Total (mW)	3.96	11.69	22.13	45.07

* 2160@120fps is equivalent in terms of throughput to 4320@30fps.



- For a fair power comparison with [6], we disregarded the power share of the blocks that are exclusive to our design:
  - Our design consumes **6.4%** and **46.5%** less power than theirs for 2160@60fps and 1080@30fps, respectively
- The work in [6] reports lower power than those in [11]–[14]

4. Conclusions

- The case study hardware architecture is more energy-efficient than state-of-the-art similar ones
- The obtained power and area breakdowns show that buffers are responsible for up to 64% of both area and power
  - The largest one (more than 30% of area and power) is intrinsic of any memory-aware parallel FME architecture
- Future work may address power optimization of buffers through low-power techniques using slower cells.

References

[1] Cisco, "VNI complete forecast highlights: Global - 2016 year in review", Cisco, Tech. Rep., Jun. 2017.

[2] ITU-T, "Rec. ITU-T H.265: High efficiency video coding", ITU, Geneva, Rec. H.265, 2013.

[3] F. Bossen *et al.*, "HEVC complexity and implementation analysis", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1685–1696, 2012.

[4] S. Blasi, I. Zupancic, E. Izquierdo, and E. Peixoto, "Adaptive precision motion estimation for HEVC coding", in *PCS'15*, 2015, pp. 144–148.

[5] C. M. Diniz *et al.*, "A reconfigurable hardware architecture for fractional pixel interpolation in high efficiency video coding", *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 34, no. 2, pp. 238–251, 2015.

[6] V. Afonso *et al.*, "Hardware implementation for the HEVC fractional motion estimation targeting real-time and low-energy", *Journal of Integrated Circuits and Systems*, vol. 11, no. 2, pp. 106–120, 2016.

[7] I. Chakrabarti *et al.*, *Motion Estimation for Video Coding: Efficient Algorithms and Architectures*, ser. Studies in Computational Intelligence. Springer International Publishing, 2015.

[8] G. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression", *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74–90, 1998.

[9] Synopsys, *Synopsys design compiler, version M-2016.SP4*. 2015.

[10] *TSMC standard cell library tcbn45gsbwptc*, TSMC, 2011.

[11] C. M. Diniz, M. Shafique, S. Bampi, and J. Henkel, "High-throughput interpolation hardware architecture with coarse-grained reconfigurable datapaths for HEVC", in *ICIP'13*, 2013, pp. 2091–2095.

[12] G. Pastuszak and M. Trochimiuk, "Architecture design of the high-throughput compensator and interpolator for the H.265/HEVC encoder", *Journal of Real-Time Image Processing*, vol. 11, no. 4, pp. 663–673, 2016.

[13] G. He, D. Zhou, Y. Li, Z. Chen, T. Zhang, and S. Goto, "High-throughput power-efficient VLSI architecture of fractional motion estimation for ultra-HD HEVC video encoding", *IEEE Trans. VLSI Syst.*, vol. 23, no. 12, pp. 3138–3142, 2015.

[14] V. Afonso *et al.*, "Memory-aware and high-throughput hardware design for the HEVC fractional motion estimation", in *SBCCI'15*, 2015, pp. 1–6.

Contact: ismaelseidel@inf.ufsc.br

50th IEEE International Conference on Circuits and Systems (ISCAS)
May 27-30, 2018– Florence, Italy