9.7 A 470mV 2.7mW Feature Extraction-Accelerator for Micro-Autonomous Vehicle Navigation in 28nm CMOS

Dongsuk Jeon, Yejoong Kim, Inhee Lee, Zhengya Zhang, David Blaauw, Dennis Sylvester

University of Michigan, Ann Arbor, MI

Recently, computer vision technologies are being applied to smaller systems such as cell phones, digital cameras, and unmanned surveillance platforms [1]. Feature extraction is a critical step in these applications. However, high-quality feature extraction algorithms require high performance and power-hungry processing, making them unsuitable for power-constrained embedded systems unless their scope is restricted based on a specific application [4]. On the other hand, algorithms with relatively low computational cost exhibit poor extraction performance and can be used only in limited application spaces, such as face detection. Hence, there is a need for high performance and energy-efficient feature extraction for use in emerging mobile applications.

This paper proposes a power-efficient speeded-up robust features (SURF) extraction accelerator targeted primarily for micro air vehicles (MAVs) with autonomous navigation (Fig. 9.7.1). Typical object recognition SoCs [4-6] employ an application-specific algorithm to choose specific regions of interest (ROIs) to reduce computation by focusing on a small portion of the image. However, this approach is not feasible in applications where the whole image must be analyzed, such as visual navigation that requires the extraction of general features to determine location or movement. In addition, multicore architectures need to run at high clock frequencies to meet high peak performance requirements and the power consumption of inter-core communication becomes prohibitive. Since feature extraction algorithms require significant memory accesses across a large area, parallelization in a multicore system requires costly high-bandwidth memories for massive intermediate data.

To overcome these challenges we propose a general full frame feature extraction hardware accelerator that can be applied to visual navigation and other applications. First, we use an accelerator-based approach to achieve high energy efficiency at low clock rates with reduced data storage and communication costs. To accomplish this, we co-optimize algorithm and architecture to obtain an efficiently parallelized FIFO-based architecture, while maintaining extraction accuracy. We also propose a description processing element with significantly reduced memory requirements and minimized data communication thanks to shared data flow. Furthermore, the low clock frequency and matched throughput of the proposed architecture improve voltage scaling headroom. Furthermore, a new shift-latch-based parallel FIFO architecture is proposed to reduce area and power.

Figure 9.7.2 shows the proposed feature extraction accelerator architecture, which operates at a low clock frequency of 27MHz with an 8b grayscale image input of one pixel per cycle. The input image is divided into 11 subsections, which are overlapped by 88 pixels to allow for feature extraction at the borders (Fig. 9.7.1). The design has two major components: detector and descriptor. The detector performs filtering and searches for local maxima in the location-scale space. The descriptor applies Haar wavelet filters to the input image and up to 40 processing elements (PEs) in parallel generate feature vectors around each interest point. Unused descriptor processors are dynamically power gated to reduce active and static power. Two identical 2D image integrators are implemented in the detector and descriptor to reduce FIFO memory space. While increasing area by 9700μm^2 over a shared integrator, this design choice reduces the FIFO memory footprint by 56% (95000μm^2) as the integrated image requires more than double the memory space compared to the original image.

Feature extraction requires large memories to store intermediate values including the integrated image, interest points, and filter responses, limiting energy efficiency. Hence, we propose a new energy-efficient shift-latch-based parallel FIFO architecture (Fig. 9.7.3). Conventional shift registers use flip-flops, which are robust at low voltage, but contain two latches per cell, and hence require nearly 2× larger power and area compared to a latch-based memory. However, latches are level-sensitive, making it impossible to shift all data within the same cycle due to race conditions. Therefore, we propose a latch-based shift register that enables only one latch at a time, allowing that latch to accept data from a neighboring input latch (Fig. 9.7.3). At this point the neighboring latch stores duplicate data, becoming a "bubble". In the next cycle, the neighboring latch is enabled and it then takes its new value from its neighboring latch. After N cycles, all values are shifted down by one entry and one output is produced from the last latch, completing one period. After N-1 periods, the value initially stored in the first latch has been shifted to the last latch and can be passed to a readout circuit, providing N(N-1) total FIFO delay with throughput of one output per N cycles. By using N identical lanes, we can obtain a one-sample-per-cycle N(N-1) entry FIFO. To avoid race conditions and improve reliability at low voltage, each lane is enabled only every other cycle in the final design. Compared to a conventional FIFO [2], the proposed shift-latch FIFO is 49% smaller with 62% lower energy and 37% speedup for a 16b 1K-entry FIFO.

In the subthreshold regime, very low on-off MOSFET current ratios lead to functional failures on shared bitlines due to large leakage current through access transistors in inactive cells. We propose a leakage balancing technique (Fig. 9.7.3, bottom) where the inactive cells are preset to have an equal number of ones and zeros, resulting in roughly balanced pull-up and pull-down leakage currents on the bitline. Since each lane is accessed sequentially, 2-transistor AND gates are employed. This technique suppresses the impact of PVT variations and improves readout delay σ by 34% with 4% speedup (despite added AND gate delay, Fig. 9.7.4).

To reduce the computational requirements of SURF [3] and target an energy-efficient, parallelized hardware implementation, we propose interest point detection using only a single-octave with an added filter size (33 pixels) and fast localization to avoid computationally extensive matrix operations while maintaining accuracy (Fig. 9.7.5). To limit the storage required by each descriptor processor, we design a circular sampling regime divided into 32 subsections for feature description. All descriptor processors simultaneously update their accumulation vectors by monitoring a single, common stream of filtered image responses, reducing each processor's memory by 89% as well as communication overhead. The results are post-processed in two efficient steps: orientation assignment based on the subsection with the largest vector magnitude (Fig. 9.7.5, lower left) and vector reordering, rotation, and normalization.

The proposed feature extraction accelerator was fabricated in 28nm LP CMOS. Measurements in Fig. 9.7.6 show that it processes 640×480 30fps input video at 470mV with a 27MHz clock frequency, compared to frequencies >100MHz for typical vision SoCs. The feature extraction core consumes 2.7mW and achieves 55.3TOPS/W, marking a 3.5× improvement over prior work (OPS/W used for comparison against other works with different functionalities). Since the proposed architecture does not vary with video size, a ~3× clock speedup allows processing of 1280×720 30fps video, consuming 12mW at 81MHz (measured at 600mV).

Acknowledgements:
Funding support by the Army Research Laboratory, IC fabrication support of STMicroelectronics and test support of Nathan Michael and Shaojie Shen at the University of Pennsylvania are gratefully acknowledged.

References:

[1] T. Kurafuji et al., "A scalable massively parallel processor for real-time image processing," *ISSCC Dig. Tech. Papers*, pp. 334-335, 2010.

[2] M. Seok et al., "A 0.27V 30MHz 17.7nJ/transform 1024-pt complex FFT core with super-pipelining," *ISSCC Dig. Tech. Papers*, pp. 342-344, 2011.

[3] H. Bay et al., "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346-359, 2008.

[4] Y.-C. Su et al., "A 52mW full HD 160-degree object viewpoint recognition SoC with visual vocabulary processor for wearable vision applications," *IEEE Symp. VLSI Circuits*, pp. 258-259, 2011.

[5] J. Oh et al., "A 320mW 342GOPS real-time moving object recognition processor for HD 720p video streams," *ISSCC Dig. Tech. Papers*, pp. 220-222, 2012.

[6] Y.-M. Tsai et al., "A 69mW 140-meter/60fps and 60-meter/300fps intelligent vision SoC for versatile automotive applications," *IEEE Symp. VLSI Circuits*, pp. 152-153, 2012.

978-1-4673-4516-3/13/$31.00 ©2013 IEEE

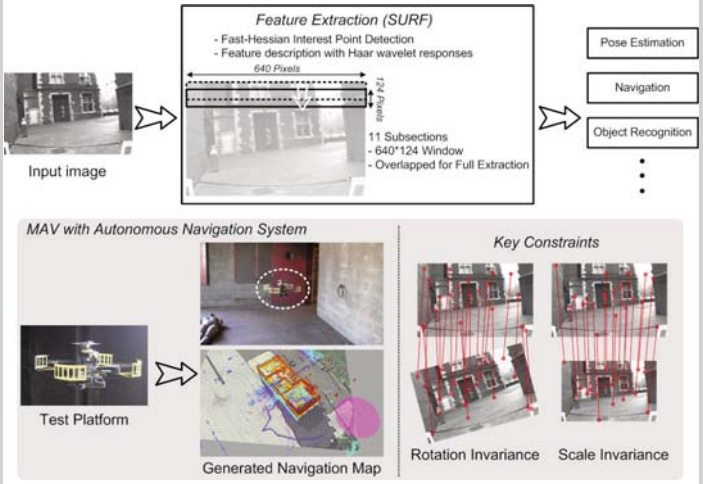


Figure 9.7.1: Applications of the proposed feature extraction accelerator (top). This accelerator is designed specifically for MAVs (Micro Air Vehicles) with autonomous navigation systems (bottom).

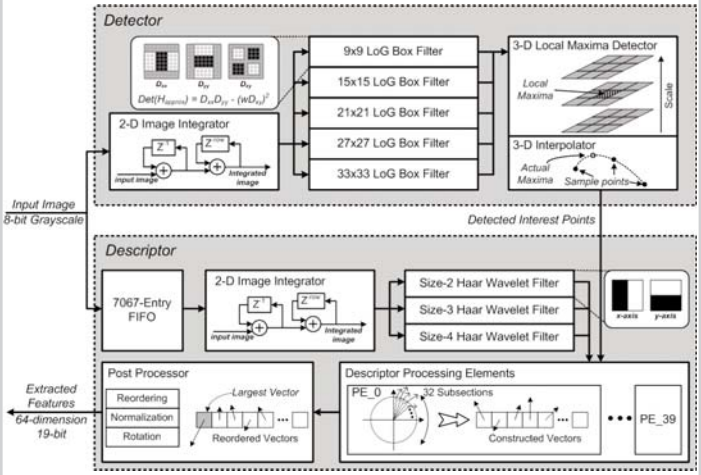


Figure 9.7.2: Proposed throughput-matched feature extraction architecture. The entire system operates at a low clock frequency that matches the input image throughput.

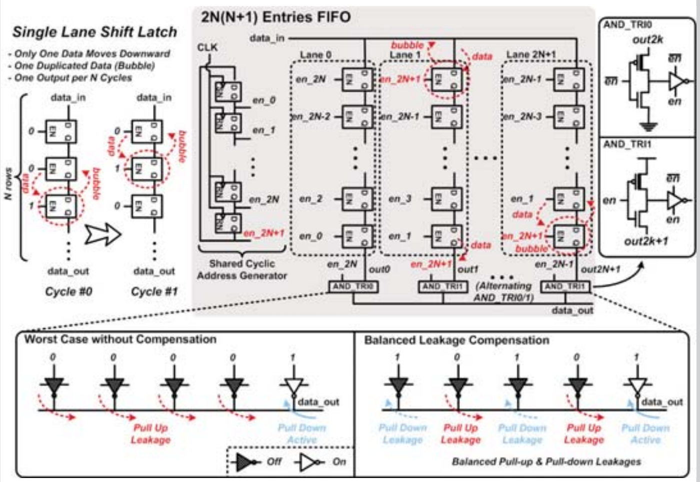


Figure 9.7.3: An energy-efficient shift-latch hybrid FIFO architecture is shown (top) along with a balanced leakage compensation technique (bottom).

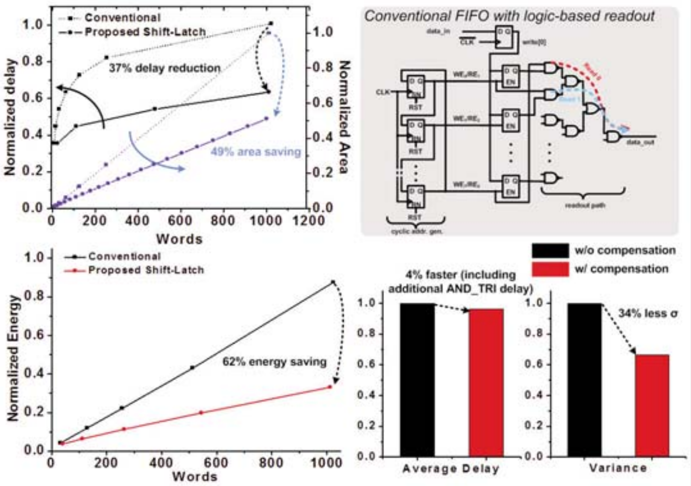


Figure 9.7.4: Simulation results of the conventional and proposed FIFO architectures.

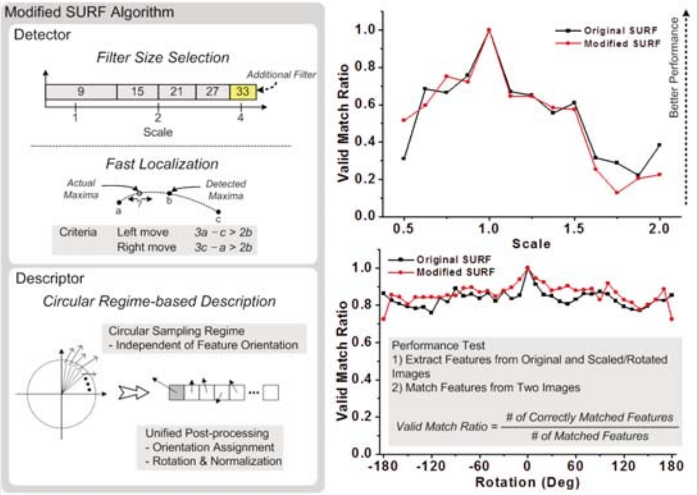
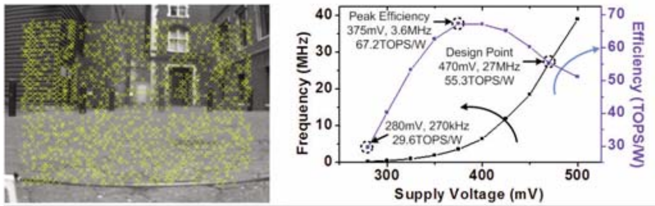


Figure 9.7.5: Algorithm-hardware co-optimization techniques (lower left). Simulations show that the optimized SURF algorithm provides similar scale- and rotation-invariance performance to original SURF (right).

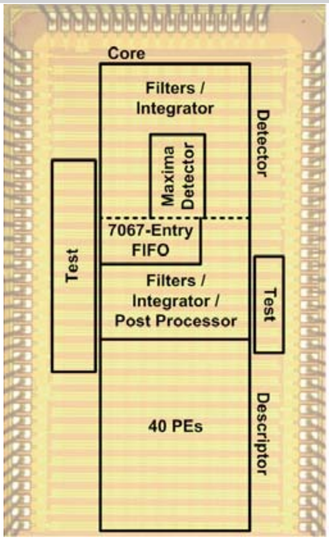


	Proposed	[4]	[5]	[6]
Technology	**28nm**	65nm	130nm	40nm
Design Target	**Feature Extraction**	Object Recognition	Object Recognition	Object Recognition
Base Algorithm	**SURF**	SIFT	SIFT	Haar-like
Extraction Scope	**Entire Frame**	ROI only	ROI only	ROI only
Input Video	**640×480**	1920×1080	1280×720	1280×960
Core Voltage	**470mV**	1V	0.7~1.2V	0.9V
Power	**2.8 mW**	52.5mW	320mW	69.3mW
Scaled Efficiency	**55.3TOPS/W***	12.4TOPS/W**	6.6TOPS/W**	15.7TOPS/W**

Scaled Efficiency = Reported Efficiency * (Technology/28nm) * (Voltage/470mV)^2

*Average efficiency with equivalent number of operations **Peak efficiency

Figure 9.7.6: A sample image is marked with 1421 features extracted using the accelerator (upper left). Measured results (upper right) and comparison table (bottom) are shown.



Technology	28nm LP CMOS
Vdd	470mV
Clock Freq.	27MHz (102FO4)
Core Area	0.85x2.61mm^2
Input Video	640x480 30fps
Power	2.7mW
Performance	149.3GOPS
Efficiency	55.3TOPS/W

Figure 9.7.7: Die photo of the feature extraction accelerator fabricated in 28nm CMOS with a summary table.