

# A Precision-Improved Processing Architecture of Physical Computing for Energy-Efficient SIFT Feature Extraction

Yi Li, Fei Qiao\*, Xinghua Yang, Qi Wei\* and Huazhong Yang\*  
Institute of Circuit and System, Dept. of Electronic Engineering, Tsinghua University  
\*Email: qiaofei@tsinghua.edu.cn

**Abstract**—A precision-improved processing architecture of physical computing for energy-efficient SIFT feature extraction algorithm has been proposed in this paper. With the novel physical computing technology of active resistor network (PC: ARN), the SIFT algorithm could be processed in analog signal domain without synchronizing clock signals, which means the complex algorithm could be completed within the setup time of the circuit. Especially for the multi-scale Gaussian convolution of SIFT algorithm, an architecture of two-layer 1-dimension PC: ARN has been adopted to compute the horizontal and vertical 1-dimension gaussian filter, in which way higher accuracy can be obtained when compared with the results processed by a 2D active circuit network. A circuit-level simulation with 65nm CMOS technology has been carried out, which shows the energy consumption of gaussian pyramid multi-scale-filtering hardware architecture is about 25.3pJ, where the size of input frame is assigned as  $256 \times 256$  pixels. Additionally, the average matching ratio in different image pairs is around 80%. Moreover, integrated into the dominating CMOS image sensor with column-parallel readout technology of analog-to-digital convertor, about 20 $\times$  speedup can be achieved comparing with previous implementations with FPGA, GPU, etc.

**Keywords**—SIFT Algorithm, Gaussian Pyramid, Physical Computing

## I. INTRODUCTION

With intelligent devices gathering more attention, new challenges on low-energy consumption and real-time requirements become severer. In conventional views, efficient visual algorithms can be achieved by designing parallel acceleration technologies like GPU and FPGA [1], [8], [9] when it is difficult to meet real-time requirements on CPU platform. The procedure of all these technologies starts from digital signals output from ADCs and then useful information are extracted as final results, during which the whole computation runs in digital system. Scale-Invariant Feature Transform (SIFT) algorithm is generally acknowledged as the most robust feature extraction algorithm in visual application. However, due to the high computational cost of multi-scale convolution, it is rather difficult to meet the real-time requirements when implementing SIFT algorithm on CPU platform. Even on GPU and FPGA with parallel acceleration technologies, it is also hard to reach real-time requirement when the resolution of input images are up to 1080p. Also, the continuous energy supply is a big challenge when considering the high energy consumption in digital system.

Since SIFT algorithm is much more robust on illumination,

rotation, and scale than any other feature extraction algorithms, more efforts have been made to accelerate SIFT algorithm in recent years. Most of these work tried to balance the tradeoff between the run-time and their computational precision in digital system. Work [1] set the parameter of Gaussian Pyramid as  $[octave, scale] = [3, 6]$ , and all the multi-scale convolution operations are implemented with parallel computing. The computational time of Gaussian Pyramid takes up 3.4ms when input image size is assigned as  $640 \times 480$ . Work [2] proposed a layer parallel SIFT with integral image to accelerate this algorithm, where the octave is set to be 2 and scale is 4. Through abundantly reducing the computational data, this work can extract about 2000 features/frame for  $1920 \times 1080$  images at 30 frames/s. To take full advantage of the data independency, GPU technologies are introduced to parallelly compute the Gaussian Pyramid. Based on GPU technology, work [8] proposed a heterogeneous dataflow scheme to make general GPU more efficient on mobile devices. In this way, it can achieve a speedup of 4-7 $\times$  over an optimized CPU version. It has a best performance on Nexus 7 platform, where it takes 28.2ms to complete Gaussian Pyramid when input image size is  $320 \times 280$ . Work [9] utilized a GPGPU platform to accelerate SIFT algorithm, which works on NVIDA cards and extracts about 800 features from  $640 \times 480$  video at 10Hz.

All the work mentioned previously are implemented in digital system essentially, where the performance is fundamentally limited by the system clock. An 2D circuit network was proposed to cope with the Gaussian Pyramid of SIFT algorithm in our previous work [3], where to accelerate the computation task we pushed the filtering operation into analog domain. To obtain higher precision based on the architecture presented in work [3], a different circuit topology is presented in this paper. Physical Computing circuit focuses on the potential mathematical properties of the devices themselves as well as the topology of circuit structure. The ideal that let physics do computes [10] is a novel approach to transfer some computations, which are much complex in digital system, to analog domain and let analog circuits complete the specific computing. In analog system, the computing procedure is free of system clock, so the whole run-time is equal to the setup time of the circuit. We will introduce the basic theory of SIFT algorithm in section II. System model and the details of Physical Computing circuit design will be presented in section III, including system framework, cascade Gaussian circuit and negative resistor. Section IV is the experimental results. Conclusion is given in section V.

## II. SIFT ALGORITHM BRIEF INTRODUCTION AND HARDWARE IMPLEMENTATION ANALYSIS

SIFT algorithm includes four stages: (1) establishing the multi-scale Gaussian Filter Pyramid and Gaussian Differential Pyramid; (2) key-points detection and refining; (3) computing and estimating the principle orientation of each key-point; (4) confirming local feature points and generating descriptors. Considering that there are more ten times data than original input image to be convolved with Gaussian function at the front-end of SIFT, most of previous work focused on the first stage and tried various methods to accelerate the algorithm. As shown in figure 1, conventional methods are demonstrated on digital platform where data are gathered through image sensors and then converted into digital signals by ADCs. 2D network data in different octaves of Gaussian Pyramid are parallelly filtered by various acceleration technologies. In our previous Work [3], ADCs are pushed backwards and Gaussian Pyramid is processed in analog domain. To balance the computational precision and hardware resource complexity, we designed three octaves and each octave comprises of six images which are convolved with six different filtered scales of Gaussian functions. The first Gaussian function scale in every octave is equal to the fourth scale of upper octave as figure 2 illustrated. For input data, we duplicate the signals to different circuit networks and all the data are filtered by different Gaussian scales.

As mentioned in previous work, the computing in each 2D network has no mutual effect to each other. So we use a 2D circuit network to complete a 2D filtering operation. All filtering operations in different octaves are parallelly executed. The processing time of a 2D circuit network depending on the RC(Resistor-Capacitance) constant parameter in each cross-node of the 2D network remains constant no matter how the scale of network is changed. Analog signal processing circuit has an ultra high processing speed, which can complete the whole convolution operations within hundreds of picoseconds. However, there are some flaws left in this work. Firstly, the connections between each nodes are very complex when the scale of the circuit becomes larger and larger. Secondly, there are still some promising work to enhance the computation precision of 2D circuit network. To resolve the two flaws mentioned, a cascade active resistor circuit network architecture constructing of two layer 1D circuit networks is proposed. In each layer of the circuit architecture the number of connections between different nodes is much less than 2D circuit network, thus we can make it more easy to arrange the wires in two separate layers.

## III. SYSTEM MODEL AND PHYSICAL COMPUTING CIRCUIT DESIGN

### A. system framework

A smart CMOS image sensor embedded with PC: ARN to perform multi-scale Gaussian convolution of SIFT algorithm is presented in this section. As described in work [3], the framework of standard commercial CMOS image sensor can be divided into two modules, photodetector and AD read-out circuits. PC: ARN is integrated between photodetector and AD read-out circuits, as figure 3 illustrates. Gaussian Pyramid circuit comprises of three octaves and there are six different

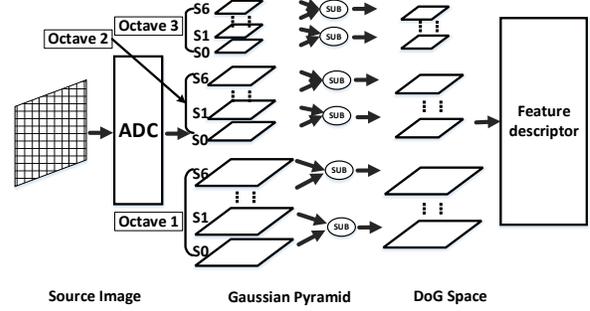


Fig. 1: Gaussian Filtering Pyramid and DoG space: input images are down-sampled as different dimensional sizes, and each size is filtered by six scale Gaussian functions [3].

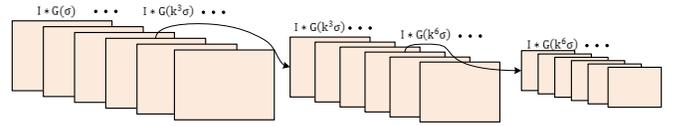


Fig. 2: Each octave comprises of six images convolved with six different scale of Gaussian functions. The first Gaussian function scale is equal to the fourth scale in upper octave.

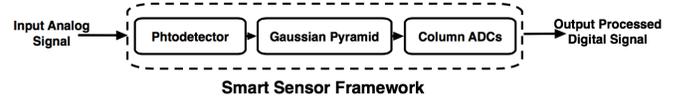


Fig. 3: Original illumination signals are detected through photodetectors, and then converted into analog voltage signals and carried to PC: ARN. Processed voltages will be converted into digital signals and readout.

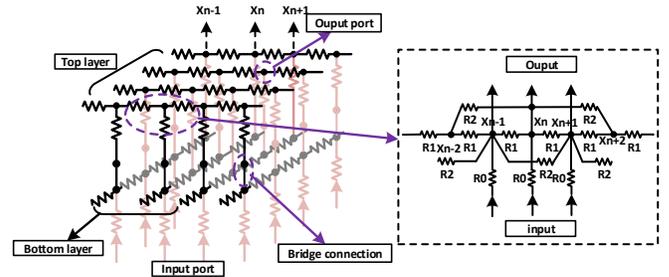


Fig. 4: left-subimage: Overview of cascade circuit architecture. The two layers are connected by a bridge connection. Every layer is fully symmetric. For one node, the connection with its adjacent nodes is shown in the right-subimage; right-subimage: 1D Gaussian circuit connection:  $R_2$ s are negative resistors,  $R_2/R_1 = -4$ ,  $R_0/R_1 = \lambda$ .

scale circuits in each octave. A cascade Gaussian filter circuit to is used to replace the 2D Gaussian filter circuit. The upper layer computes the vertical convolution, and then the next layer circuits computes the horizontal convolution.

### B. Two-layer 1-dimension Gaussian Filtering Circuit

In visual applications, the convolution operation of a 2D Gaussian function and an image is usually transformed into

two cascade 1D Gaussian functions convolving this 2D image. Assign the image as  $I(x, y)$ , and 2D Gaussian function as  $G(x, y)$ , then

$$\begin{aligned}
 G(x, y) * I(x, y) &= \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} G(k, l) I(x - k, y - l) \\
 &= \frac{1}{2\sigma^2} \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} e^{-\frac{k^2+l^2}{2\sigma^2}} I(x - k, y - l) \\
 &= \frac{1}{2\sigma^2} \sum_{k=0}^{m-1} e^{-\frac{k^2}{2\sigma^2}} \left\{ \sum_{l=0}^{n-1} e^{-\frac{l^2}{2\sigma^2}} I(x - k, y - l) \right\}
 \end{aligned} \tag{1}$$

Where we assume the width and height of the Gaussian filter as  $m$  and  $n$  respectively, and  $\sigma$  is the scale of Gaussian function.

Based on the inherent property that a 2D Gaussian function can be decomposed to the convolution result of two 1D Gaussian functions, as presented in Eq.(1), a two-layer circuit is designed to filter a 2D image. The Gaussian scale is controlled by the resistor  $R_0$  value in this circuit. As figure 4 shows, a bridge connection is used to connect the upper layer with the bottom layer circuit network. The left sub-figure in figure 4 gives the overview of two-layer circuit architecture.  $x_{n-1}, x_n, x_{n+1}$  can be replaced by each other, that to say all nodes in this network are totally symmetric. Taking node  $x_n$  as the central one, the connection with the adjacent nodes is illustrated as the right sub-figure in figure 4. Resistors in this circuit can be grouped as three categories.  $R_1$  and  $R_2$  should have the specific proportional  $R_2/R_1 = -4$ . On the condition that we assign a variable  $\lambda$  as  $\lambda = R_0/R_1$ , then the Gaussian scale can be adjusted by modifying the value of  $\lambda$  [5], [6]. Input ports are connected with one side of  $R_0$ , and the other side of  $R_0$  connects with output ports. Bridge connections combine the output ports of upper layer with input ports of the second layer.

### C. Negative Resistor

As illustrated in figure 4,  $R_2$  is assigned to be a floating negative resistor, which should be symmetric at both ports. To realize its symmetric property, two negative resistors are put together to compose a double-end resistor. A negative resistor comprises of two Op-Amps and six resistors, shown in figure 5. The equivalent resistance value can be yielded,

$$R_1 = -\frac{R_{11}}{R_{12}} R_{13}, R_2 = -\frac{R_{21}}{R_{22}} R_{23} \tag{2}$$

To guarantee the equation  $R_1 = R_2$ , then

$$\frac{R_{11}}{R_{21}} = \frac{R_{12}}{R_{22}}, R_{13} = R_{23} \tag{3}$$

Eq.(2) and Eq.(3) indicate that the left-end equivalent resistance value  $R_1$  is equal to the right-end resistance value  $R_2$ . In this way, we can take this module as an ordinary resistor.

## IV. EXPERIMENTAL RESULTS

### A. Simulation Environment Setup

The circuit-level simulation runs on APS tool embedded in Cadence platform. The supply voltage is set as 1V. The size

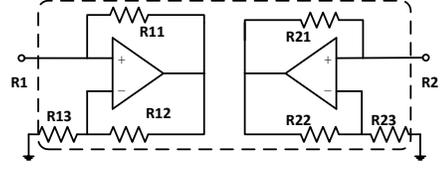


Fig. 5: negative resistor

TABLE I: scale values

octave	1	2	2.5	3.2	4	5.1
octave 1	1.6	2	2.5	3.2	4	5.1
octave 2	3.2	4	5.1	6.4	8.1	10.2
octave 3	6.4	8.1	10.2	12.9	16.2	20.4

of input image is customized as  $256 \times 256$ . The value of  $R_1$  is set to be  $400\Omega$ , and  $R_2$  is  $-1k\Omega$ . Since the Gaussian Pyramid has twelve different scale values, the values of  $R_0$  are variable in different circuits and the different values of  $\sigma$  are listed in table I. To evaluate the performance of the whole Hybrid system, column read-out ADCs with 500Ms/s maximal pixel rate [4] are introduced to convert analog signals into digital bits.

### B. System Function Precision

We know that the basic kernel function of SIFT algorithm is Gaussian function, so it is fundamentally important to accurately simulate Gaussian function with an analog circuit. In this work, abundant data about the system function of cascade Gaussian filter circuit have been compared with the 2D Gaussian convolution circuit proposed in work [3]. Due to the peak error/pixel in work [3] is up to 20%, it will badly influence the system function precision. Considering the property of 2D Gaussian function, we try to use two-layer-1D Gaussian circuit to complete convolution task, so that we can obtain more accurate Gaussian function curve.

We collected the relative error data of the circuit system function in work [3] and we found that the maximal error appears at the minimal Gaussian scale. With the value of Gaussian scale growing, the peak error/pixel becomes smaller. As recorded in table II, the relative error generated from 2D circuit is less than 10% when Gaussian scale is higher than 4. The main factor to introduce so big deviation is the asymmetric bleeder circuit path in 2D circuit at every node. To avoid the side effect of bleeder circuit, cascade circuit architecture is

TABLE II: System Function Precision with Different values of  $\sigma$ . Average error=total error/the number of pixels, and maximal error is the maximal single-pixel error among all the pixels in a circuit network when comparing with ideal gaussian function.

$\sigma$	this work		work [3]	
	average error	maximal error in pixel	average error	maximal error in pixel
1.6	1.28%	5.65%	6.95%	20.47%
2	1.07%	4.43%	6.82%	19.89%
2.5	0.81%	3.14%	5.21%	16.78%
3.2	0.78%	2.06%	4.78%	12.62%
4	0.72%	1.38%	4.52%	10.19%
5.1	0.66%	1.90%	4.19%	7.10%
6.4	0.62%	1.85%	4.04%	5.01%
8.1	1.23%	2.35%	3.90%	4.41%
10.2	1.22%	2.53%	3.76%	4.32%
12.9	1.14%	2.51%	3.62%	4.54%
16.2	1.16%	2.58%	2.90%	4.95%
20.4	1.14%	2.56%	2.09%	4.55%

TABLE III: Performance and Energy Consumption on Various Platform

Technology	Configuration	Run Time	Energy Consumption
CPU technology [1]	2.1GHz Core	443ms	29.12J
GPGPU on general platform [9]	NVIDIA 9000GTX	5.05ms	0.5J
GPU on Mobile device [8]	NVIDIA Tegra 3	20.6ms	2.06mJ
FPGA technology [1]	Altera DE2-70	0.73ms+0.064ms(ADC time)	292 $\mu$ J
Physical Computing circuit [3]	2-D Gaussian Pyramid circuit	138.03ps+0.504ms(ADC time)	74.79pJ
Physical Computing circuit(this work)	two-layer 1-D Gaussian Pyramid circuit	100ps+0.04ms(ADC time)	25.3pJ

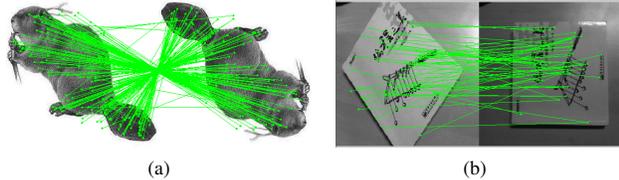


Fig. 6: Feature points matching results: There are generally more than 80% feature points matched. In every feature point pair, there are two to three pixels position offset.

introduced, as described in section III. The peak error/pixel of the system function of the novel circuit is less than 6%, shown in table II. Compared with the circuit proposed in work [3], we obtain about 4 $\times$  harvest in precision when using this novel circuit architecture.

### C. Performance and Energy Consumption

To evaluate the circuit setup time, a general parasitic capacitance is introduced. As described in introduction section, SIFT algorithm has been implemented on various platforms. For instance, work [1] use FPGA to accelerate SIFT. In such way, by reducing some computational data and improve the parallelism in computing, it can realize 900 $\times$  speedup when comparing with conventional CPU technologies. Work [9] proposed a general GPU to accelerate SIFT, which utilizes the parallel computing in GPU to realize acceleration purpose. As presented in table III, Physical Computing circuit can process the complex computing in 100ps, which is much faster than previous methods. Considering the interface to digital system, we introduce a general column-readout technology and the ADC rate is 500MS/s, so the AD conversion time is up to 0.04ms, which becomes the bottleneck of the whole hybrid system.

We have packaged the Gaussian Pyramid circuit as a Physical Computing kernel to replace the multi-scale convolution in SIFT algorithm. With the simulation of APS, matlab and gcc platforms, we implement the whole SIFT algorithm. For  $256 \times 256$  pixel images, 70% ~ 80% feature points can be detected and matched correctly in several groups tests. In figure 6(a), on the situation that only rotation operation is taken between the two images, more than 80% points are successfully matched. And in figure 6(b), more than 70% points are matched correctly when rotation, illumination and scale are changed.

## V. CONCLUSION

In this paper, a cascade Gaussian circuit architecture is presented to realize the complex computation in SIFT algorithm.

Different from conventional analog circuit, Physical Computing circuit focuses on the inherent mathematical property of electronic devices and the circuit topologies. Since Physical Computing circuit runs in analog domain, it is completely free of system clock. All the process time of the whole procedure equals to the setup-time of the analog circuit, which could be completed within 1ns, so it can obtain  $10^6 \times$  speedup when comparing with FPGA and any other digital platforms. The whole Physical Computing circuit is integrated into a CMOS image sensor, and all data output are multi-scale convolution results. Due to the computation of processing Gaussian Pyramid is pushed to the front-end sensor, a large amount energy consumption will be saved. We know smart devices is a hot topic these years, so low-energy and high-speed smart CMOS image sensor is a tendency in vision application. By integrating Physical Computing circuit into CMOS image sensor, we move some computational operations into front-end to ease the burden of the subsequent digital system. Considering that ADCs will be introduced as the connection of Physical Computing circuits and subsequent digital system, the ADCs performance will become the bottleneck of the hybrid system. So how to reduce the data amount to be converted into digital bits is another challenge in our design.

## REFERENCES

- [1] Huang, Feng-Cheng, et al. "High-performance SIFT hardware accelerator for real-time image feature extraction." *Circuit and Systems for Video Technology, IEEE Transactions on* 22.3(2012): 340-351.
- [2] Chiu, Liang-Chi, et al. "Fast SIFT design for real-time visual feature extraction" *Image Processing, IEEE Transactions on* 22.8 (2013): 3158-3167.
- [3] Yi Li, Fei Qiao et al, "Physical Computing Circuit With No Clock to Establish Gaussian Pyramid of SIFT Algorithm." *international symposium on circuits and system ISCAS 2015* for printed.
- [4] Lim, Seunghyun, et al. "A 240-frames/s 2.1-Mpixel CMOS image sensor with column-shared cyclic ADCs." *Solid-State Circuits, IEEE Journal of* 46.9 (2011): 2073-2083.
- [5] White, Joseph L., and A. A. Abidi, "Active resistor network as 2D sampled data filter," *Circuit and System I: Fundamental Theory and Applications, IEEE Transactions on* 39.9(1992): 724-733.
- [6] Kobayashi, Haruo, Joseph L. White, and Asad A. Abidi, "An active resistor network for Gaussian filtering of images," *Solid-State Circuits, IEEE Journal of* 16.5(1991): 738-748.
- [7] Lowe, David G, "Object recognition from local scale-invariant features." *Computer vision, 1999. The processing of the seventh IEEE international conference on*, Vol. 2. IEEE, 1999.
- [8] Rister, Blaine, et al. "A fast and efficient SIFT detector using the mobile GPU" *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013.
- [9] Sinha, Sudipta N., et al. "GPU-based video feature tracking and matching" *EDGE, Workshop on Edge Computing Using New Commodity Architectures*. Vol. 278. 2006.
- [10] Chandramoorthy N, Swaminathan K, Cotter M, et al. "understanding the landscape of accelerators for vision[C]" *Signal Processing Systems (SIPS), 2014 IEEE Workshop on*. IEEE 2014: 1-6.