INTEGRATION, the VLSI journal 47 (2014) 417–430

Contents lists available at ScienceDirect

INTEGRATION, the VLSI journal

journal homepage: www.elsevier.com/locate/vlsi

A digital array based bit serial processor for arbitrary window size kernel convolution in vision sensors

Mehdi Habibi *, Alireza Bafandeh, Muhammad Ali Montazerolghaem

Department of Electrical Engineering, University of Isfahan, Isfahan, Iran

ARTICLE INFO

Article history:
Received 5 February 2013
Received in revised form
29 November 2013
Accepted 29 November 2013
Available online 7 December 2013

Keywords:
Array-based
Bit serial
Vision sensors
Kernel convolution
Digital processing

ABSTRACT

The high speed and in-pixel processing of image data in smart vision sensors is an important solution for real time machine vision tasks. Diverse architectures have been presented for array based kernel convolution processing, many of which use analog processing elements to save space. In this paper a digital array based bit serial architecture is presented to perform certain image filtering tasks in the digital domain and hence gain higher accuracies than the analog methods. The presented method benefits from more diverse convolution options such as arbitrary size kernel windows, compared with the digital pulse based approaches. The proposed digital cell structure is compact enough to fit inside an image sensor pixel. When incorporated in a vision chip, resolutions of up to 12 bit accuracy can be obtained in kernel convolution functions with $35 \times 28\ \mu m^2$ layout area usage per pixel in a 90 nm technology. Still, higher accuracies can be obtained with larger pixels. The power consumption of the approach is approximately 10 nW/pixel at a frame rate of 1 kfps.

1. Introduction

With the ever increasing demand for complex vision systems, the use of powerful processors with higher processing speed and lower power dissipation is necessary. In-pixel digital processing solutions are advantageous compared with external digital image processors for two main reasons; the processing speed can be increased significantly and the dynamic power dissipation can be reduced considerably. The reduction of dynamic power is possible since with multiple in-pixel digital processors, the clock pulse frequency can be reduced. Subsequently with a lower clock frequency the supply voltage can be lowered, resulting in overall reduction of dynamic power dissipation [1].

CMOS image sensors have the ability to perform all or part of the required processing inside each pixel using active MOSFET devices [2]. These types of image sensors with processing capabilities are known as vision chips. Since image processing and especially machine vision tasks are diverse, usually low level and fundamental tasks are implemented in the sensor structure. Kernel convolution is one of the most important low level image processing functions used in machine vision tasks. It is effective in noise removal, edge detection, image sharpness and softness adjustment, directional filters, image compression and many other applications.

Although different vision chips have been presented that perform processing using analog photodiode data [3], however analog processing circuits usually lack accuracy since the current and voltage parameters typically change between 1–5% of the full swing value depending on the device geometry [4]. Thus most vision chips which perform some type of in-pixel signal processing in the analog domain usually have a pixel accuracy of approximately 7–8 bits [5,6].

If the pixel data is prepared in binary form, the digital binary processing of data can be accomplished without the drawbacks of the analog approach. To this day, many image sensors with in-pixel ADC structures have been presented to provide high speed and accurate data output [7–9]. Furthermore, using multi-tier chip technology, the combination of a digital pixel sensor and FPGA die can result in programmable vision sensors without sacrificing the pixel fill factor [10–12]. An important issue is that it is difficult for analog circuits to benefit from device scaling due to device mismatches and subsequently accuracy issues. However the device scaling is advantageous in the digital approaches since it reduces the circuit size and allows the processor to be integrated inside each pixel [13,14].

It should be noted that digital pulse based schemes have been presented that are appropriate and effective for dedicated image processing tasks such as cellular neural vision systems [15]. The sensors presented in [16,17] can perform addition and multiplication directly on pulse trains by using pulse addition or pulse division (reduction). The direct processing of pulse trains suffers from the drawback that the convolution options are relatively limited and kernel convolution window sizes of more than 3×3

* Corresponding author.
E-mail address: mhabibi@eng.ui.ac.ir (M. Habibi).

0167-9260/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.vlsi.2013.11.007

pixels is relatively cumbersome to implement. This is a limiting factor in adaptive machine vision tasks where the kernel size needs to be adjusted over a wide dynamic range.

In this paper a kernel convolution vision chip is presented which can process image data and perform desired filtering functions in real time. The processing is performed in the digital

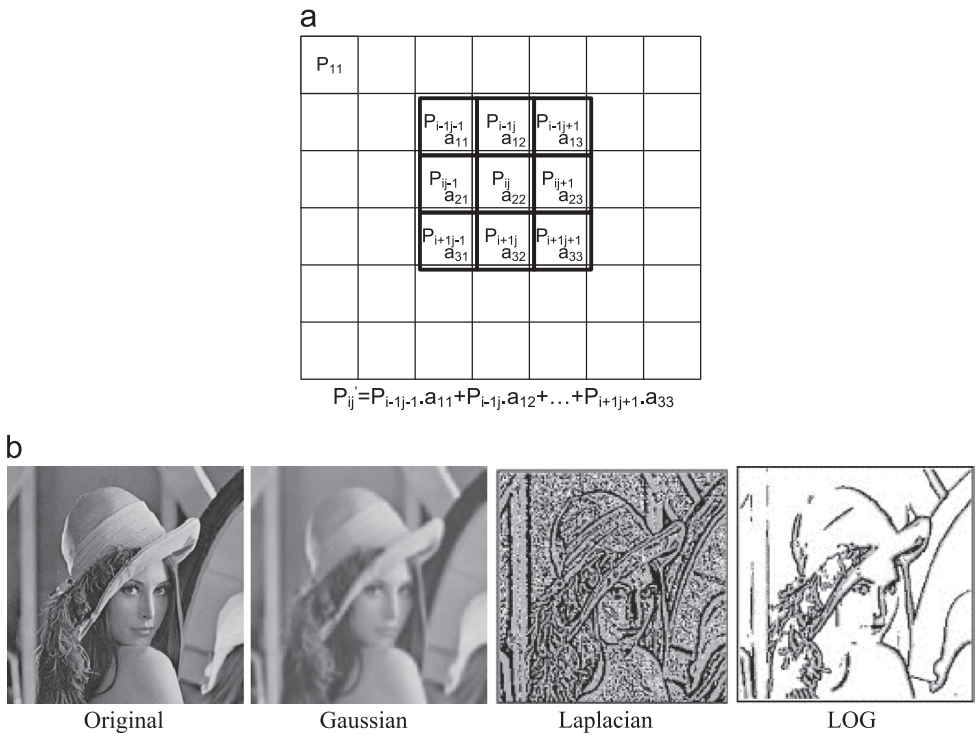

Fig. 1. (a) Basic kernel convolution principle. (b) Different image filtering functions implemented by different convolution kernels.

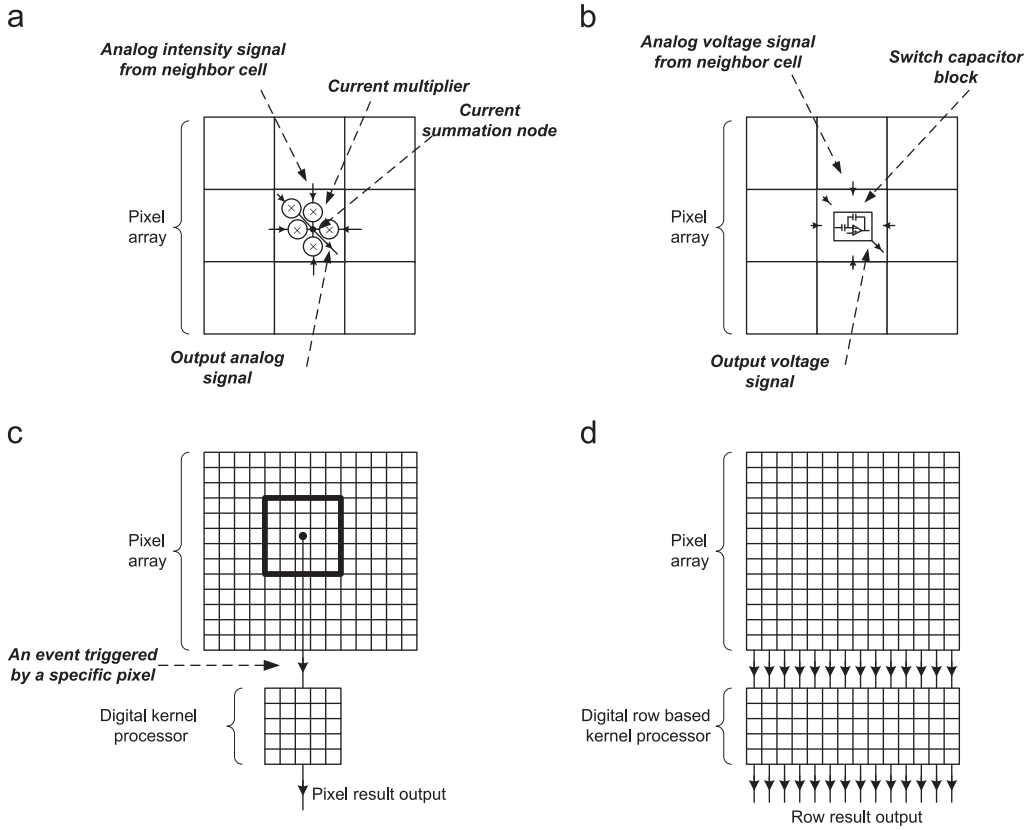

Fig. 2. Different hardware solutions used for kernel convolution. The processor is based on (a) Pixel based analog multipliers, (b) Pixel based switch capacitor blocks, (c) Digital event based computation block and (d) Digital row based computation block.