System-level design space identification for Many-Core Vision Processors

Jones Yudi, Carlos Humberto Llanos, Michael Huebner

1. Introduction

The evolution of the microelectronics technology has enabled the industry to integrate billions of transistors on a single chip, following the well-known Moore’s Law prediction. This evolution enabled the development of new ideas, products and whole new markets based on the embedded processing systems. Following this trend, some new concepts emerged some years ago: the Cyber-Physical Systems (CPS), the Ubiquitous Computing (UC) and the Internet of Things (IoT), which are expected to change the interaction between humans and the surrounding environment. The computing capabilities embedded in different devices, daily-life objects, buildings, and so on, will be distributed, ubiquitous and transparent to the users [1].

Distributed embedded devices and intelligent sensors will collect and process huge amounts of data, enabling the pervasiveness of the computational environment [2]. These sensors, also known as Smart Sensors, are devices able to perform not only pre-processing algorithms but also more complex applications with embedded and stand-alone intelligence. The Smart Cameras, a particular type of Smart Sensor, will be essential as devices able to capture and interpret the human behavior in an environment, among other events/objects. These future cameras must be able to perform complex applications simultaneously, coping with real-time constraints. Also, other important requisites must be fulfilled: energy consumption, chip temperature control, reliability, Quality of Service (QoS), data security, privacy, power management, cost (silicon area), and so on [3].

Image Processing and Computer Vision (IP/CV) applications are computationally costly, mainly due to the huge amount of data to be processed. New technologies are enabling the use of increasing resolutions, and new applications are demanding high-performance with tight deadlines. In this scenario, there are cam-
eras with frame rates over one hundred thousand frames per second in high resolutions. When the application allows storing the images for later processing, the problem is then related to the amount of storage memory. However, in the context of real-time applications, the tight timing constraints can lead the system to failure when not attended. The success of Real-Time IP/CV processing systems is highly dependent on how efficiently the inherent parallelism is explored and, therefore, a special hardware/software architecture must be used to fulfill the application’s needs [4].

Simple processing architectures, such as the General-Purpose Processors (GPPs), are not able to provide a power efficient real-time performance for IP/CV applications. The sequential nature of GPPs does not offer many opportunities for parallelism exploration, which means that high operating frequency is necessary, implying in more power dissipation. In this context, there are single-core alternative architectures, such as the VLIW, which can overcome simple RISC-based GPPs by exploring more efficiently the Instruction-Level Parallelism (ILP) for IP/CV applications [5]. However, the VLIW architectures do not scale well for a high number of issues, mainly due to the register file and multiplexing overhead, limiting the amount of processing parallelism [6].

For several years, to enhance the processor’s speed, the industry relayed on increasing the transistor count per area, and the circuit switching frequency. However, problems with the power density (related to the Dennard Scaling model) limited this trend [7]. The industry entered then in the so-called “post-Dennardian Scaling Era” when it was forced into a transition to multi-core systems [8]. On the other hand, this architectural change was already expected, due to the natural evolution of parallel computing area from multi-computer clusters to intra-chip multi-processors, which are envisioned to provide more scalable power efficiency when increasing the number of cores [9]. To be able to handle all the constraints and also offer enough flexibility for different IP/CV applications, a solution for the Smart Cameras is also to migrate from single to multi/many-core processing architectures [10].

A deep analysis of the hardware characteristics is essential to overcome the limitations of Smart Camera devices, from image acquisition aspects to processing architectures features. From a historical perspective, the first commercial cameras used Charge-Coupled Devices (CCDs), which dominated the market for some decades. Several technological aspects, such as manufacturing difficulties and power consumption, limited the use of CMOS technology for image sensors. However, in the 1990’s, the CMOS-APS invention allowed for an efficient integration of all image sensor electronic components in a single chip [11]. The advances in the CMOS technology (mainly due to the processor’s market needs), were also used by the camera sensors manufacturers, enabling huge cost reduction, and replacing the CCD cameras for the CMOS ones in most devices.

Fig. 1 shows different sensor acquisition configurations [3], which provide single or multiple pixel streams. Each configuration has its advantages and disadvantages, implying a trade-off analysis by the system designer. The Single Pixel configuration is the most common and offers the smallest cost. However, it presents also a bottleneck similar to single-port memory: only a single pixel stream appears to the processing part (not shown in the picture). The other three configurations offer more parallelism, but also extra silicon area costs. The One per Pixel configuration provides full pixel access parallelism, however, it is not cost effective due to the area overhead, and the severe fill factor reduction [12]. The Parallel Columns and Region-based configurations offer a balance among silicon area, fill factor, and acquisition parallelism. Real-Time IP/CV systems should start the processing phase as soon as the pixels are available. The acquisition architecture choice depends not only on the speed and parallelism but also based on the IP/CV algorithms to be executed. A quantitative analysis of the acquisition schemes is provided as part of the DSE in Section 3.2.

Two similar architectural concepts, Focal-Plane Image Processing (FPIP) and Near-Sensor Image Processing (NSIP), have emerged in the late 1980’s and early 1990’s, as a solution for high-speed acquisition and processing in Real-Time IP/CV systems. The basic idea was to merge the pixel sensors with processing elements in the camera’s focal-plane [13], or as close as possible one to the other [14], with the goal of offering pixel access parallelism and low latency. Fig. 2 illustrates the idea [15,16,10]. Several architectures from the literature use this concept. Most of them were implemented completely with analog circuits (both acquisition and processing) [14,17,18]. The analog implementations have advantages over digital ones, such as power consumption and speed. However, they also have a lack of programmability, and their implementation complexity makes them hardly adopted in IP/CV applications. Also, due to physical limitations, the scale of integration achieved better results for digital circuits than for analog ones [19]. Digital technology offers the advantages of flexibility through programmability, the reuse of standard cells and more scalability, in comparison with the analog circuits. Due to these aspects, modern implementations unify analog and digital circuits as follows: analog ones for the acquisition and ADC, and digital ones for the processing part [20,21,22,23].

Along the years, analog/digital integration achieved success. However, most implementations are quite simple from the processing architecture point of view. Most architectures were more focused on showing the feasibility of the concept through prototyping, without exploring the huge design space available for both acquisition and processing architectures. A clear trend is the in-

---

**Fig. 1.** CMOS sensor configurations: single and multiple pixel streams [3].
دریافت فوری متن کامل مقاله

<table>
<thead>
<tr>
<th>ISI Articles</th>
<th>مرجع مقالات تخصصی ایران</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ امکان دانلود نسخه تمام متن مقالات انگلیسی</td>
<td>✓ امکان دانلود نسخه ترجمه شده مقالات</td>
</tr>
<tr>
<td>✓ پذیرش سفارش ترجمه تخصصی</td>
<td>✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله</td>
</tr>
<tr>
<td>✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله</td>
<td>✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب</td>
</tr>
<tr>
<td>✓ دانلود فوری مقاله پس از پرداخت آنلاین</td>
<td>✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات</td>
</tr>
</tbody>
</table>