Low-Power Vision Chips based on Focal-Plane Feature Extraction for Visually-Assisted Autonomous Navigation

Ricardo Carmona-Galán

Abstract— Avoiding obstacles and finding the way around are tasks that can greatly benefit from an efficient implementation of vision. While higher level vision can be performed by conventional microprocessors at an acceptable rate, lower level vision represents a heavy computational load to deal with. The usual sensor plus ADC plus microprocessor scheme either fails to meet the timing requirements or fails to operate under a low power budget. Since the information contained in the visual stimulus is highly redundant, converting every single pixel value to digital prior to any processing is inefficient. Instead, we are working in adapted architectures in which the parallelism that is inherent to lower level vision tasks is largely exploited. This hierarchical approach emulates the organizational principles of biological vision systems, by using an array of elementary and relatively coarse processors to achieve global computation, and also the operation of the elementary cells, by using analog and mixed-signal processing building blocks. Our chips are capable of efficiently extracting image features and salient points at the focal plane in order to facilitate the task of identifying objects and interpreting the scene.

I. INTRODUCTION

The implementation of vision in autonomous robotic platforms represents a challenge in which efficiency in the computation is an unavoidable request [1]. In contrast with other application fields, like digital still photography or consumer electronic cameras, it is not sufficient to obtain a high quality picture. The outcome of the required processing is not an image, but a description of the scene with valuable information to assist independent navigation. Vision is understood as a cognitive procedure by which the closest environment is perceived and interpreted. Its implementation requires a master plan oriented towards perception [2]. In this context, the traditional approach to image processing [3], which is not different from the conventional digital signal processing scheme in which sensing is immediately followed by digitization, memory and serialized processing, does not pass the efficiency test. The conventional approach, though universal, does not match the multidimensional structure of images and video sequences. Moreover, it does not take advantage of the inherent redundancy of the visual stimulus, in which the value of a single isolated pixel is not as important as the aggregation of pixels. Therefore, the conventional image processing scheme in which every individual pixel value is converted to digital prior to any processing is inefficient when applied to the implementation of vision. An alternative approach is to bring part of the processing elements and memory closer to the sensor array. In one

of the aspects, the implementation of distributed computing and memory resources, this is what lies behind multi-core processors [4] [5]. By splitting heavy tasks in parallel threads the overall performance can be boosted without increasing power consumption [6]. This is achieved by exploiting parallelism in the data and algorithms structure. Highly parallel architectures trade flexibility for an extraordinary increase in computing efficiency. In addition to the advantage of using an adapted architecture, i. e. an architecture that matches the inherent parallelism of early vision tasks —in which the same operation needs to be performed on each and every pixel and most of the time with independence on what is happening beyond its closest neighborhood—, the redundancy existing in the visual stimulus permits a moderate accuracy requirement (6-7 bits) for in-pixel processing elements. Analog circuits can be employed to implement some of the processing blocks in a efficient manner [7] [8]. The highest figures are found for architectures with a reduced programmability, i. e. highly optimized for a specific task, for instance Gaussian filtering [9]. The performance of a system based on this type of devices is improved not only because of the energy efficiency of the focal-plane processing. As the output of the vision chip is a simplified representation of the scene being surveyed, with less data but of a higher abstraction level, the subsequent digital processor does not have to cope with the heavy computational loads associated to low-level vision. Clock frequency and memory access can be thus significantly reduced.

In this review paper we try to illustrate this hierarchical approach to vision with some examples. First we will consider the speed and power requirements on a vision chip in the field of robot vision. Then we will display the characteristics of a bio-inspired hierarchical vision processing scheme aiming to energy efficiency. Then we will consider a couple of examples, developed in our labs in which focal-plane processing provides the support for the efficient generation of an information-enriched representation of the scene on a single chip.

II. ENERGY EFFICIENCY IN VISION CHIPS

Let us consider an simplified model for image processing and vision. Any vision algorithm is composed of a number of tasks that start from an input image and terminate with either another image, a different type of representation of the scene, a warning signal, a trigger, etc. In all cases, a number of elementary operations, N_{op}, is performed to the original image pixels and the derived intermediate data. In this simplified model, all the elementary operations require

the same amount of energy to be performed, let us say e_0, and take the same amount of time to be completed, t_0. Power consumption of an elementary processor computing just this operation will be e_0/t_0. The total amount of energy required to complete the transformation of the imput image into its corresponding output will be:

$$E_{\text{tot}} = N_{\text{op}} \cdot e_0 \qquad (1)$$

and the total amount of time required to complete the algorithm is:

$$T_{\text{tot}} = \frac{N_{\text{op}}}{N_{\text{proc}}} \cdot t_0 \qquad (2)$$

where N_{proc} is the number of elementary processors that will be operating in parallel. This is an oversimplification given that algorithms can not be parallelized at wish and the speedup achieved by parallelization is limited, as stated by Amdahl's law [6]. However, this simplified model permits analyzing the specifications of a vision system for different scenarios. First, consider the case of time-critical applications. In this context, it is important to minimize the time it takes for the algorithm to be executed, Eq. (2), or equivalently to maximize the system speed:

$$\text{Speed} = \frac{1}{T_{\text{tot}}} = \frac{N_{\text{proc}}}{N_{\text{op}}t_0} \qquad (3)$$

Then, in order to meet very tight timing requirements, we need to either increase the number of processors, N_{proc}, or reduce the time it takes for each elementary processor to complete a single operation, t_0; and/or reduce the total number of operations required to complete the algorithm, N_{op}. This last alternative can be implemented by approximation [10], image simplification at early stages [11], sparse representation of the scene [12] or hierarchical processing and resource optimization [13].

For systems running on a low power budget, energy efficiency is required, therefore power consumption needs to be minimized:

$$\text{Power} = \frac{E_{\text{tot}}}{T_{\text{tot}}} = N_{\text{proc}} \frac{e_0}{t_0} \qquad (4)$$

what is done by lowering the number of processors that are working in parallel, N_{proc}, or using low-power elementary processors.

Finally, for time-critical and low-power applications, that are usual in fields like robotics, unmanned vehicle navigation, autonomous surveillance camera nodes, etc., the figure of merit to be maximized is:

$$\text{FoM} = \frac{\text{Speed}}{\text{Power}} = \frac{1}{E_{\text{tot}}} = \frac{1}{N_{\text{op}}e_0} \qquad (5)$$

Here, the effect of parallelization is not represented by an explicit factor, like N_{proc} in Eqs. (3) and (4). However, parallelization has an effect in this FoM, given that distributed resources require less power to interact and do it much faster. As in Eq. (3), modifying the algorithm in order to reduce the number of operations to be realized increases the FoM. In addition, using more efficient elementary processors, i. e.

reducing the amount of energy that needs to be invested on each elementary operation, contributes to maximize the FoM in Eq. (5) as well. e_0 is commonly measured in nJ/OP (nanojoules per operation). For a fixed throughput, which is usually the case in systems in which the data rate is determined by the nature of the incoming real-time signal, e. g. speech or video, increasing the energy efficiency is equivalent to reducing the power dissipation for the prescribed number of operations per second, what is usually measured in MOPS/mW [14]. Reducing e_0 can be done at two levels. At system level, using parallel processing architectures with distributed sensing, processing and memory [15] eliminates the time and power overheads dedicated to data access and communication across the system —and this is where the degree of parallelism implicitly affects the FoM; or at circuit level, using analog circuits that render a highly efficient implementation for a moderate accuracy.

III. HIERARCHICAL VISION PROCESSING

The different tasks that constitute the visual processing chain can be separated into a hierarchy defined by the abstraction level (Fig. 1). In fact, the amount of data to be handled at each stage observes an inverse relation with the complexity of the data structures. For the earliest vision stages, input data are the raw readings of the sensors. The number of data to be processed is high, but they are simple in their internal structure and the process flow can be very regular. At the other end of the processing chain, those tasks of a higher cognitive nature are usually executed on a reduced set of data structures, this time of a highly symbolic content. This functional hierarchy can be ideally implemented in a physical structure composed of various layers (tiers). This arrangement permits a progressive reduction of the data structures to be processed while maintaining a fully parallel connection across tiers. Coincidentally, a similar scheme is found in biological vision systems [16], in which vertical interactions across a layered structure [17] and massively parallel processing [18] have been documented.

The major obstacle to the implementation of this architecture in planar technologies is the trade-off between image
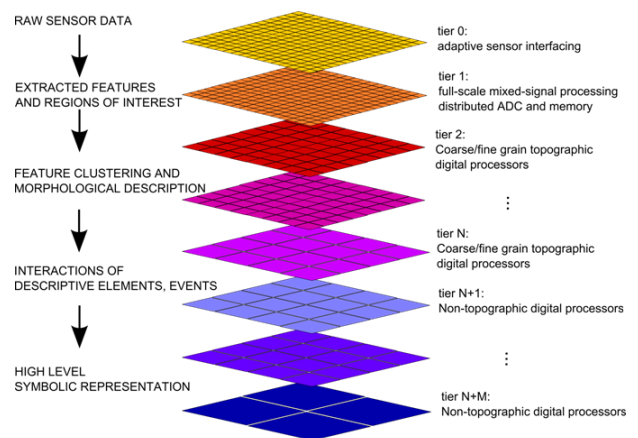

Fig. 1. Hierarchical organization of vision tasks

size and processing speed. On one side, incorporating some processing elements to the focal plane increases the size of the elementary processor, thus influencing image resolution, fill factor and ultimately the effective image size. On the other side, if the sensor size is optimized by eliminating in-pixel processing, data transmission bottlenecks occur between sensor, memory and processor, thus compromising the processing speed. One alternative to overcome this limitation, this is to have a high degree of parallelization while maintaining a reasonable pixel pitch, can be 3D circuit integration [19].

IV. Chip Examples

Our approach to the implementation of these organizational principles is supported by two elements: conveying the heavy computational tasks that can be parallelized close to the sensors and using power efficient analog and mixed-signal circuits to implement the elementary in-pixel processing blocks. In the first example, Gaussian filtering at the focal-plane can help reducing the amount of data to be handled by subsequent digital processing. In the second, extended functionalities like DoG, minima and maxima detection and fully parallel image digitization is implemented in a vertical integration technology.

A. Scale-space generator on-a-chip

Gaussian kernels are a fundamental component of a computational approach to visual perception motivated by physics and biological vision [20]. Convolution with Gaussian kernels and Gaussian derivatives constitute a canonical class of image operators for early vision. They are able to generate a scale space [13] and, consequently, a multi-scale image representation [21] of the scene. It is worth mentioning that scale-space operators have a similar form to the receptive fields observed in neuro-physiological studies [22]. This type of image representation is certainly useful for image interpretation. As there is no a priori knowledge about the scale of the relevant elements in the scene, a multi-scale representation covers all the possible ranges. Image features can then be extracted at different scales and scale-invariant features can be highlighted as characteristic of whatever takes
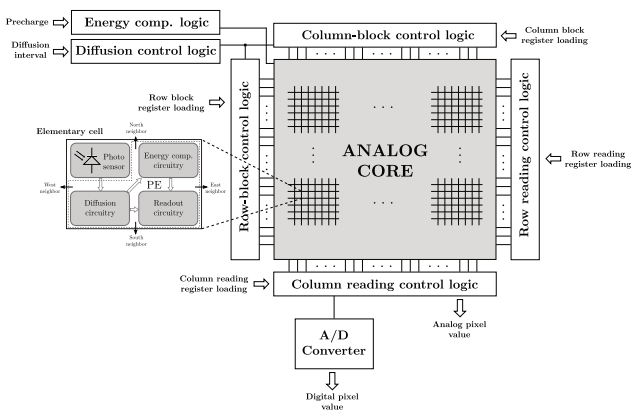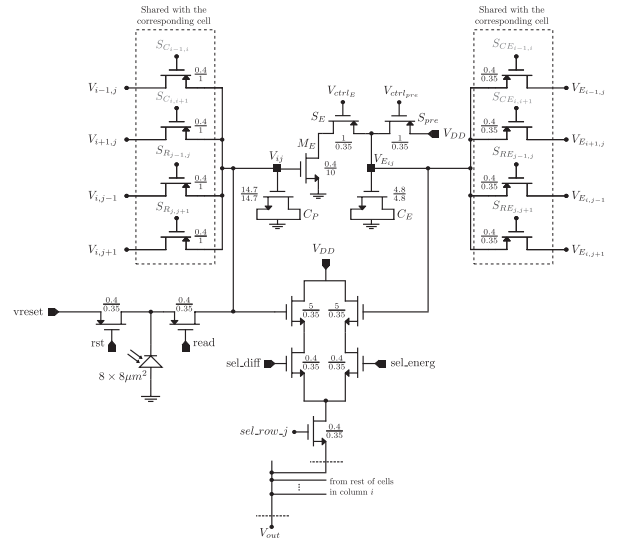

Fig. 3. Elementary cell of the array

place in the visual field [23]. It is not strange that visual attention models based on saliency make extensive use of these operators [24].

A prototype chip implementing on-chip scale-space generation is reported in [9]. It is a 176×144-px smart image sensor which implements a massively parallel SIMD-based focal-plane processing array composed of pixel-level processing elements (PE). These PEs, which carry out analog image processing concurrently with photosensing, can be grouped into fully-programmable rectangular-shape areas by loading the appropriate interconnection patterns into the registers at the edge of the array. The targeted processing can be thus performed block-wise. The architecture of the chip is depicted in Fig. 2. The power consumption associated to the capture, processing and A/D conversion of an image flow at 30fps, with full-frame processing but reduced frame size output, ranges from 2.7mW to 5.6mW, depending on the operation to be performed. The chip has been designed and fabricated in a 0.35μm CMOS-OPTO process. The chip contains around half million transistors, 98% of them
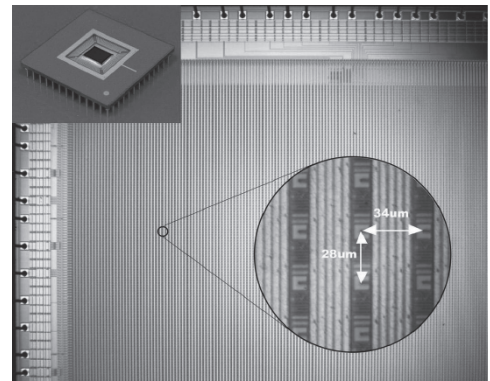

Fig. 2. Floorplan of the prototype chip

Fig. 4. General view and microphotographs of the prototype
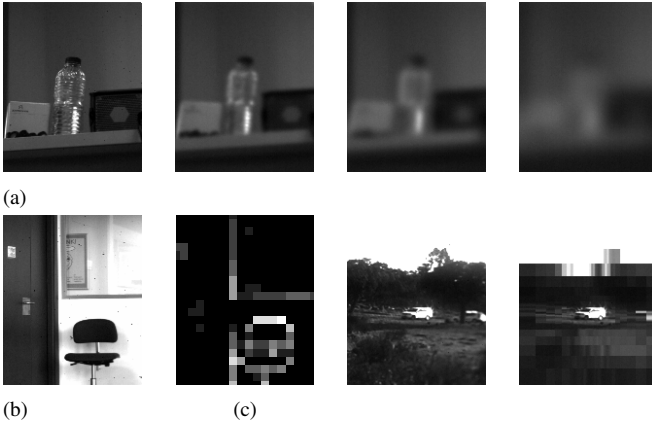
Fig. 5. Output images: (a) scale-space and (b) energy-based representations and (c) foveated image

Fig. 6. Layout of the mixed-signal processor array

working in analog mode.

The schematic of the elementary cell of the analog core is depicted in Fig. 3. Each PE contains a photosensor and a state capacitor, C_P, that is 4-connected to its neighbors through p-type MOS transistors. The equivalent resistance R_{eq} of these transistors along with the value of C_P determine the time constant $\tau = R_{eq}C_P$ of the resulting MOS-based RC network, which is intended to implement Gaussian filtering by time-controlled diffusion [25]. The scale parameter ξ is related with width of the filter $\xi = \sigma^2$, which in turn is determined by the ratio between the time interval in which the network is permitted to evolve and the time constant of the network: $\sigma = \sqrt{2t/\tau}$. In addition to this operation, that can be selectively applied to different sub-images, it is possible to generate an energy-based representation of the scene. This is useful for the efficient segmentation of spatially-repetitive patterns and dynamic textures. It also accounts for the amount of contrast associated with an image block thus allowing for a first estimation of the salient regions of the scene [26].

The main characteristics of the chip are summarized in Table I. A general view of the packaged prototype is shown in Fig. 4 along with a microphotograph of the chip with a close-up of the photosensors. Fig. 5 displays several output images rendered by the chip, consisting in the scale-space of the scene, the energy-based representation and a foveated

TABLE I
SUMMARY OF THE PROTOTYPE CHIP FEATURES

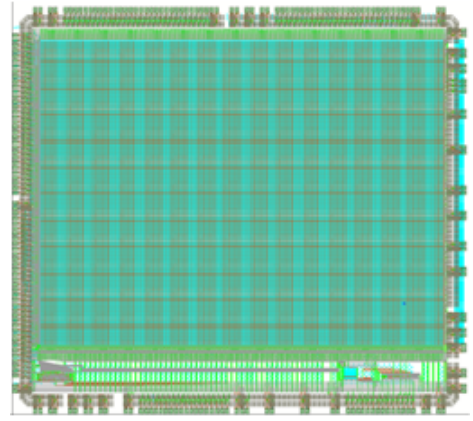|	
Technology	$0.35\mu m$ CMOS 2P4M
Die size (with pads)	$7280.8\mu m \times 5780.8\mu m$
Cell size	$34.07\mu m \times 29.13\mu m$
Fill factor	6.45%
Resolution	QCIF: 176×144 px
Power supply	3.3V
FPN	0.72%
Sensitivity	0.15V/(lux·s)
Measured power consumption (worst case)	5.6mW@30fps 22×18px
ADC throughput	0.11MSa/s ($9\mu s$/Sa)

frame. A measure of its power efficiency is given by the fact that the chip can compute a Gaussian filtered version of the image using 20nJ. Being a 176×144-pixel array, its equivalent computing power is 1.27e+6MOPS/mW. This is an impressive number, but the level of programmability of this array processor is greatly reduced. Its functionality is to realize Gaussian filters.

B. Analog front-end for a vision system on-a-3D-chip

Another example of how massively-parallel analog and mixed-signal array processing enables low-power vision implementation can be found in the design of the VIS-CUBE [27]. The target application of the VISCUBE chip is UAV navigation and exploration. Hence, the algorithmic requirements are intended to perform moving platform video analytics, including feature point extraction, displacement calculation —optical flow—, and analysis in multiple windows [28]. Image capturing and processing speed needs to reach 1000fps. In order to fulfill these requirements, our design supports multi-scale and multi-fovea processing. The common feature of these techniques is that they significantly reduce the amount of data to be processed. On one side, multi-scale algorithms handle subsampled versions of the image; hence part of the processing is done over lower resolution images. On the other side, foveated processing techniques apply early image processing techniques for the entire image, and detailed image analysis only on small windows of the original image. During a pre-processing phase, the regions of interest (ROI) are detected. Windows containing these ROIs are then cut out and further analyzed
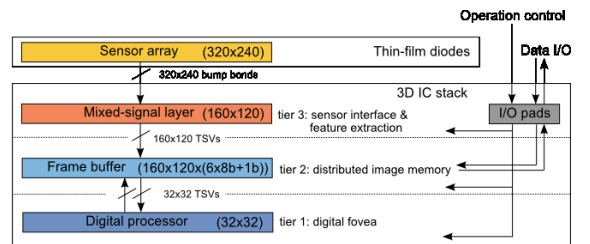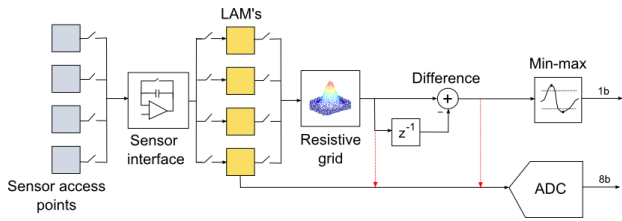

Fig. 7. General scheme of the 3D chip

Fig. 8. Conceptual schematics of the elementary mixed-signal processor

#	Description	operations	Total
8	Amplifier reset	1 sample + 1 hold	16
2	Min-max detection	256 comparisons + 256 AND	1024
6	A/D conversions	256 comparisons	1636
1	Binning	4 additions + 1 scaling	5
40	Diffusion cycles	(2 samp. + 2 hold)*4neighbors	640
	TOTAL		3321

Fig. 9. Variation of the average power consumption of tier-3 with operating temperature

in detail. The size and the resolution —or scale— of these windows depend on the type of analysis applied. The design approach for this 3D vision chip is completely different compared to the design of traditional planar chips. When selecting image processor devices we need to study the efficiency of different architectures. For early vision tasks, mixed-signal locally interconnected processor arrays and pipelined digital processor arrays can be efficiently applied [29]. Above 1000fps and under low latency requirements in near-sensor processing, a fine-grain mixed-signal processor array is the best choice. For fovea-type post processing, coarse-grain digital processor arrays can provide efficient solution. In this way, two different cellular arrays fulfill the computational requirements for the target application. The fine-grain processor (tier-3 in Fig. 7) contains the sensor, the sensor interface, mixed-signal and logic processing elements, A/D converter and analog and logic distributed memories. The foveal processor (tier-1 in Fig. 7) is built from an array of locally interconnected digital processors and their local memory. To efficiently connect cellular processor arrays with different resolutions, we need an intermediate element to perform switching, windowing and preparing data for the multi-scale foveal processor array. This is implemented by a dual-port frame buffer (tier-2 in Fig. 7). This architecture is mapped on the layers provided by a vertical integration technology [30]. The backside-illuminated sensor array, built in a thinned semiconductor layer, is bump-bonded on top of the 3D IC stack, what provides close to 100% fill factor. This practice permits using different types of sensor materials, what defines the wavelength of the radiation to be detected. This can range from visible to NIR, or even SWIR. The topmost layer of the stack (Fig. 6) contains a mixed-signal processor array responsible for sensor interfacing, low-level vision tasks and feature extraction.

The sensor interface, which is part of the mixed-signal front-end (Fig. 8), consists in a capacitive transimpedance amplifier (CTIA). It has been employed traditionally to interface arrays of passive pixels, however, it is used nowadays in high precision readout ICs [31], because several reasons: high programmable charge-to-voltage conversion ratio; low readout noise; photodiodes kept at a programmable bias voltage, that is constant along the integration time; sensing capacitance is isolated from diode parasitics. All of these characteristics make it a universal front-end for photodiodes of different materials. Although the photosensor pitch is 15μm, given by the minimum distance at which bonding bumps can be deposited, a larger silicon area is required

to allocate enough processing at the focal plane. For this reason, each PE is interfaced to 4 photodiodes. This means a 30μm pitch and a 160×120-PE array. In addition to the sensor interface, each PE contains (Fig. 8): 4 local analog memories (LAM); a tree of switches that allows computing the difference and the average of the memory contents and permits scaling, auto-zeroing and correlated-double sampling (CDS) [32]; part of a switched-capacitor network employed to realize a discrete-time emulation of the linear diffusion of the pixel values [33]; and a comparator [34] that is part of the in-pixel A/D converter and of the extrema locator circuit. The operation of this circuit consists in: capturing a 320×240-pixel frame, converting it to digital format; averaging 4-pixel groups and down-sampling the original image to 160×120 pixels; then performing 2 different Gaussian filters, converting them to digital; computing the DoG and converting it to digital; finding the local maxima and minima of this DoG and elaborating binary images containing the extrema locations. In order to do this with minimum hardware, some of the analog processing blocks in this layer have been re-used by time multiplexing. In particular, there is a pipeline that employs the two amplifiers to realize different functions at each stage.

Due to the complexity of the 3D integrated circuit structure, the different tiers have been simulated alone. These simulations have been done after parasitic extraction, and have been realized at the different technology corners. Mismatch between equally designed devices due to process parame-

ter scattering have been also contemplated. An interesting result is the average power consumption of the whole tier-3 array, for the different technology corners, in an interval of temperatures that ranges from 0C to 110C (plotted in Fig. 9). For estimating the number of operations per second realized by the mixed-signal layer (tier-3), consider that all the operations listed in Table II take place in 1ms. The result is 3.32GOPS per PE. As we have 160×120, for the complete tier-3 we have 63.74e+3GOPS, what divided by the 50mW average consumption, makes 1274.9MOPS/mW.

V. Conclusions

Focal-plane processing is a promising alternative to the efficient implementation of vision in applications fields in which low-power and high-speed are a must, like navigation in autonomous robotic platforms. Planar technologies impose a trade-off between processing speed and image resolution. Vertical (3D) integration technologies can help to overcome these limitations. From the point of view of the system design, a hierarchical approach to vision tasks can help optimize the use of the resources and therefore contribute to an energy efficient implementation.

References

[1] M. Schweitzer, H. J. Wuensche, "Efficient keypoint matching for robot vision using GPUs". *IEEE 12th International Conference on Computer Vision Workshops*, pp. 808-815, Sept.-Oct. 2009.

[2] A. C. Bovik, "Perceptual Video Processing: Seeing the Future," *Proceedings of the IEEE*, Vol. 98, pp. 1799-1803, 2010.

[3] R. Gonzalez, R. E. Woods, *Digital image processing*. (3rd Ed.), Prentice-Hall, 2008.

[4] *Intel Atom processor Z5XXA*. Tech. Rep. 319535-003US, Intel Corp. June 2010.

[5] *Tegra product roadmap*. Tech. rep., Nvidia Corp. 2010.

[6] M. D. Hill, M.R. Marty, "Amdahl's Law in the Multicore Era". *IEEE Computer*, Vol. 41, No. 7, pp. 33-38, July 2008.

[7] P. Dudek and P. Hicks, "A general-purpose processor-per-pixel analog SIMD vision chip". *IEEE Trans. on Circuits and Systems-I* Vol. 52, No. 1, pp. 13-20, 2005.

[8] M. Gottardi, N. Massari and S. Jawed, "A 100uW 128x64 pixels contrast-based asynchronous binary vision sensor for sensor networks applications". *IEEE Journal of Solid-State Circuits* Vol. 44, No. 5, pp. 1582-1592, 2009.

[9] J. Fernández-Berni, R. Carmona-Galán and L. Carranza González, "FLIP-Q: A QCIF Resolution Focal-Plane Array for Low-Power Image Processing". *IEEE Journal of Solid-State Circuits*, Vol. 46, No. 3, pp. 669-680, March 2011.

[10] M. Aubury and W. Luk, "Binomial filters". *The Journal of VLSI Signal Processing*, Vol. 12, No. 1, pp. 35-50, 1996.

[11] N. Balakrishnan et al. "A new image representation algorithm inspired by image submodality models, redundancy reduction, and learning in biological vision". *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 27, No. 9, pp. 1367-1378, Sept. 2005.

[12] J. Romberg, "Imaging via Compressive Sampling". *IEEE Signal Processing Magazine*, Vol. 25, No. 2, pp. 14-20, March 2008.

[13] T. Lindeberg, "Scale-Space", in *Wiley Encyclopedia of Computer Science and Engineering*, Wiley, 2007.

[14] T. D. Burd and R. W. Brodersen, "Processor design for portable systems". *The Journal of VLSI Signal Processsing*, Vol. 13, No. 2, pp. 203-211, 1996.

[15] E. Herrero, J. Gonzalez and R. Canal, "Distributed cooperative caching: an energy efficient memory scheme for chip multiprocessors". *IEEE Trans. on Parallel and Distributed Systems*, Vol. 23, No. 5, pp. 853-861, May 2012.

[16] R. H. Masland, "The fundamental plan of the retina". *Nature Neuroscience*, Vol. 4, No. 9, pp. 877-886, 2001.

[17] B. Roska and F. Werblin, "Vertical interactions across ten parallel, stacked representations in the mammalian retina". *Nature*, Vol. 410, pp. 583-587, 2001.

[18] B. Roska, A. Molnar, and F. S. Werblin, "Parallel Processing in Retinal Ganglion Cells: How Integration of Space-Time Patterns of Excitation and Inhibition Form the Spiking Output". *Journal of Neurophysiology*, Vol. 95, pp. 3810-3822, June 2006.

[19] P. Garrou, C. Bower, P. Ramm, *Handbook of 3D Integration: Technology and Applications of 3D Integrated Circuits*, Wiley-VCH, 2008.

[20] B. t. H. Romeny, *Front-End Vision and Multi-Scale Image Analysis*, Springer, 2003.

[21] B. Jahne, "Multiresolutional signal representation" in *Handbook of Computer Vision and Applications*, Vol. 2, pp. 67-90, Academic Press, 1999.

[22] R. E. Soodak, "Two-dimensional modeling of visual receptive fields using Gaussian subunits". *Proc. of the National Academy of Sciences*, Vol. 20, pp. 9259-9263, December 1986.

[23] D. G. Lowe, "Object recognition from local scale-invariant features". *Proc. of the IEEE Int. Conference on Computer Vision*, Vol. 2, pp. 1150-1157, 1999.

[24] L. Itti, C. Koch, E. Niebur, "A model of saliency-based visual attention for rapid scene analysis". *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 20, pp. 1254-1259, November 1998.

[25] J. Fernández-Berni and R. Carmona-Galán, "All-MOS Implementation of RC Networks for Time-Controlled Gaussian Spatial Filtering". *Int. J. Circuit Theory and Applications*, Vol. 40, No. 8, pp. 859-876, Aug. 2012.

[26] Y. Ni, "Smart image sensing in CMOS technology". *IEE Proceedings–Circuits, Devices and Systems*, Vol. 152, No. 5, pp. 547-555, 2005.

[27] P. Foldesy et al. "3D multi-layer vision architecture for surveillance and reconnaissance applications". *Proc. 19th European Conf. on Circuit Theory and Design* (ECCTD'09), pp. 185-188, Antalya, Turkey, August 2009.

[28] A. Zarandy et al. "Displacement calculation algorithm on a heterogeneous multi-layer cellular sensor processor array". *Proc. 12th Int. W. Cellular Nanoscale Networks and Their Apps*. (CNNA 2010), pp. 171-176, Berkeley, California, February 2010.

[29] A. Zarandy and Csaba Rekeczky, "2D operators on topographic and non-topographic architectures —implementation, efficiency analysis, and architecture selection methodology". *Int. J. of Circuit Theory and Applications*, Vol. 39, No. 10, pp. 983-1005, Oct. 2011.

[30] *MITLL Low-Power FDSOI CMOS Process Design Guide*. Revision 2008:6, Sept. 2008.

[31] J. N. Helou et al. "0.18um CMOS fully differential CTIA for a 32x16 ROIC for 3D ladar imaging systems". *Proc.of SPIEs Infrared and Photoelectronic Imagers and Detector Devices II*, Vol. 6294, pp. 09-13, 2006.

[32] C. C. Enz, and G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization". *Proceedings of the IEEE*, Vol. 84, No. 11, pp. 1584-1614, 1996.

[33] M. Suárez et al., "Switched-Capacitor Networks for Scale-Space Generation". *20th European Conference on Circuit Theory and Design* (ECCTD 2011), pp. 189-192, Linkoping, Sweden, August 2011.

[34] M. Suárez-Cambre et al. "Offset-Compensated Comparator with Full-Input Range in 150nm FDSOI CMOS-3D Technology". *First IEEE Latin American Symposium on Circuits and Systems* (LASCAS 2010), Iguazu Falls, Brasil, February 2010.