BOTTLENECK PROBLEM SOLUTION USING BIOLOGICAL MODELS OF ATTENTION IN HIGH RESOLUTION TRACKING SENSORS

Alexander Fish, Orly Yadid-Pecht

Abstract: Every high resolution imaging system suffers from the bottleneck problem. This problem relates to the huge amount of data transmission from the sensor array to a digital signal processing (DSP) and to bottleneck in performance, caused by the requirement to process a large amount of information in parallel. The same problem exists in biological vision systems, where the information, sensed by many millions of receptors should be transmitted and processed in real time. Models, describing the bottleneck problem solutions in biological systems fall in the field of visual attention. This paper presents the bottleneck problem existing in imagers used for real time salient target tracking and proposes a simple solution by employing models of attention, found in biological systems. The bottleneck problem in imaging systems is presented, the existing models of visual attention are discussed and the architecture of the proposed imager is shown.

Keywords: Bottleneck problem, image processing, tracking imager, models of attention

ACM Classification Keywords: B.7.0 Integrated circuits: General, I.4.8 Image processing and computer vision: scene analysis: tracking

1. Introduction

Driven by the demands of commercial, consumer, space and security applications, image sensors became a very hot topic and a major category of high-volume semiconductor production [1]. This is due to the imminent introduction of imaging devices in high volume consumer applications such as cell phones, automobiles, PC-based video applications, "smart" toys and, of course, digital still and video cameras. While most of consumer applications are satisfied with relatively low-resolution imagers, image sensors used for target tracking in space, navigation and security applications require high spatial resolution. In addition, these tracking sensors are usually supposed to provide real time tracking after multiple targets in the field of view (FOV), such as stars, missiles and others. The demands for high resolution and real time performance result in a bottleneck problem relating to the large amount of information transmission from the imager to the digital signal processing (DSP) or processor and in bottleneck in performance, caused by the requirement to process a large amount of information in parallel. The simple solution to the bottleneck in performance is to use more advanced processors to implement the required tracking algorithms or to use dedicated hardware built specially for the required algorithm implementation [2]. However, the solution to the performance bottleneck still does not relax the requirement for the large data transmission between sensor and processor.

The same problem exists in biological vision systems. Compared to the state-of-the-art artificial imaging systems, having about twenty millions sensors, the human eye has more than one hundred million receptors (rods and cones). Thus, the question is how the biological vision systems succeed to transmit and to process such a large amount of information in real time? The answer is that to cope with potential overload, the brain is equipped with a variety of attentional mechanisms [3]. These mechanisms have two important functions: (a) attention can be used to select relevant information and/or to ignore the irrelevant or interfering information; (b) attention can modulate or enhance the selected information according to the state and goals of the perceiver. Numerous research efforts in physiology were triggered during the last five decades to understand the attention mechanism [4-12]. Generally, works related to physiological analysis of the human attention system can be divided into two main groups: those that present a spatial (spotlight) model for visual attention [4-6] and those following object-based attention [7-12]. The main difference between these models is that the object-based theory is based on the assumption that attention is referenced to a target or perceptual groups in the visual field, while the spotlight theory indicates that attention selects a place at which to enhance the efficiency of information processing.

The design of efficient real time tracking systems mostly depends on deep understanding of the model of visual attention [12-14]. This paper briefly describes spotlight and object-based models of attention and proposes a solution for the bottleneck problem in image systems for salient targets tracking based on the study and utilization of the spatial (spotlight) model of attention. Two possible sensor architectures are presented and discussed.

Section 2 briefly describes the bottleneck problem in high resolution imaging systems. A review of existing models of attention is presented in Section 3. Section 4 presents descriptions of two sensor architectures, comparing it with the existing spatial (spotlight) models of attention. Section 5 concludes the paper.

2. The Bottleneck Problem in High Resolution Image Systems

As mentioned in section 1, two bottlenecks exist in high-resolution image systems: the data transmission bottleneck and performance bottleneck. The solution for the bottleneck in performance relates to increasing the processing power. This can be performed in the following ways: (a) to use more advanced processors to implement the algorithms, (b) to use dedicated hardware built especially for the required algorithms implementation and (c) employing more than one DSP/processor. Although these solutions are expensive and can dramatically increase the cost of the whole system, they provide simple and trustworthy solutions. Fig. 1 shows an example of such a kind of an imaging system. As can be seen, the system consists of the image sensors array (can either be implemented as a Charge Coupled Device (CCD) or as a standard CMOS imager, as will be described below), a number of processors (or DSPs) and a memory.

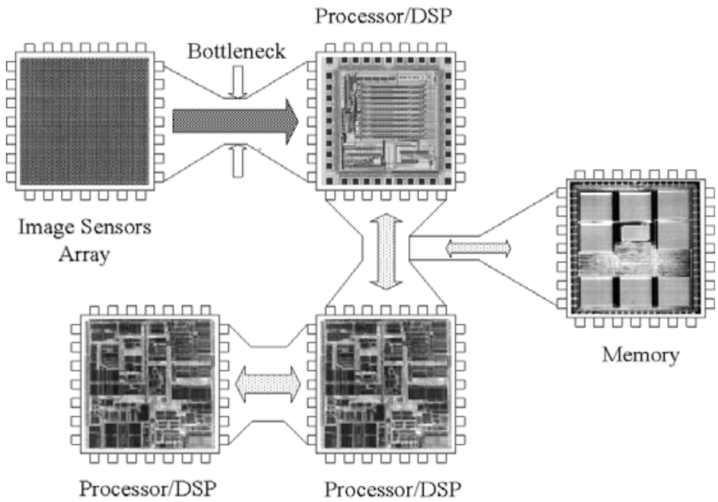

Fig.1. An example of a typical image system, incorporating an image sensor array, processors/DSPs for image processing and memory

The data transmission bottleneck from the image sensors array to the processor/DSP (see Fig. 1) is more difficult for solution. The most efficient way to solve the problem is on-chip implementation of some algorithms that can relax the information bottleneck by reducing the amount of data for transmission. However, this solution is almost impossible in CCD sensors, which cannot easily be integrated with standard CMOS analog and digital circuits due to additional fabrication complexity and increased cost. On the other hand, CMOS technology provides the possibility for integrating imaging and image processing algorithms functions onto a single chip, creating so called "smart" image sensors. Unlike CCD image sensors, CMOS imagers use digital memory style readout, using row decoders and column amplifiers. This readout allows random access to pixels so that selective readout of windows of interest is allowed. In this paper all further discussions will be related to CMOS image sensors.

Fig. 2 shows an example of an imaging system employing a "smart" CMOS image sensor and a single processor/DSP. As can be seen, in this system the image processing can be performed at three different levels: (a) at the pixel level – CMOS technology allows insertion of additional circuitry into the pixel, (b) on-chip image processing and, of course (c) image processing by processor/DSP. The larger amount of processing performed in the first two levels, the less amount of information necessary to be transmitted from the CMOS imager to the DSP. In addition, smaller computation resources are required.

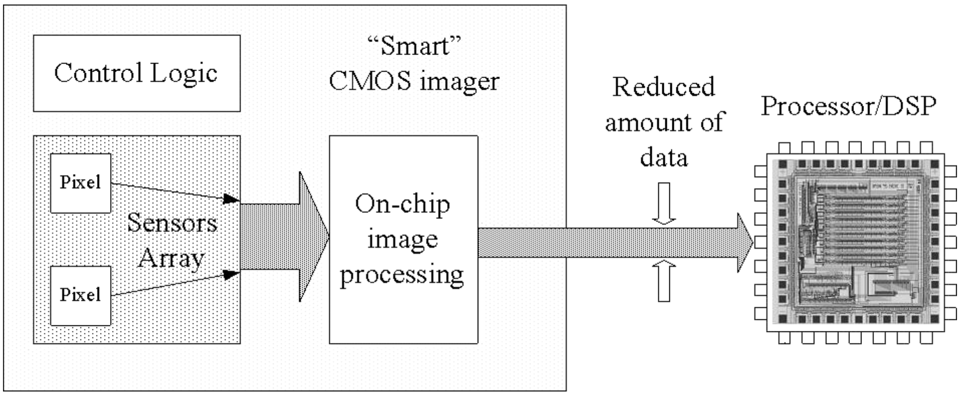

Fig.2. An example of an imaging system, employing a "smart" CMOS image sensor with on-chip processing and processors/DSPs for image processing

Generally, image processing at early stages (pixel level and on-chip level) can solve both data transmission bottleneck and performance bottleneck problems. Tracking imager architectures, proposed in this paper, will try to imitate the models of attention emphasizing on the requirement to perform the most of processing at early stages.

3. Visual Models of Attention

Although many research efforts were triggered during the last decades and numerous models of attention have been proposed over the years, there is still much confusion as to the nature and role of attention. Generally, two models of attention exist: spatial (spotlight) or early attention and object-based, or late attention. While the object-based theory suggests that the visual world is parsed into objects or perceptual groups, the spatial (spotlight) model purports that attention is directed to unparsed regions of space. Experimental research provides some degree of support to both models of attention. While both models are useful in understanding the processing of visual information, the spotlight model suffers from more drawbacks than the object-based model. However, the spotlight model is simpler and can be more useful for tracking imager implementations, as will be shown below.

3.1 The Spatial (Spotlight) Model

The model of spotlight visual attention mainly grew out of the application of information theory developed by Shannon. In electronic systems, similar to physiological, the amount of the incoming information is limited by the system resources.

Fig.3 (a). An example of spatial filtering

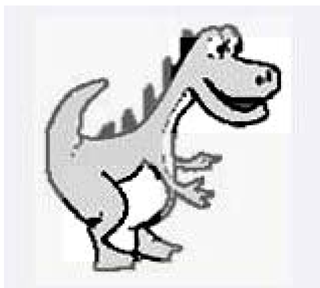

Fig.3 (b). An example of spotlight model of attention

There are two main models of spotlight attention. The simplest model can be looked upon as a spatial filter, where what falls outside the attentional spotlight is assumed not to be processed. In the second model, the spotlight serves to concentrate attentional resources to a particular region in space, thus enhancing processing at that location and almost eliminating processing of the unattended regions. The main difference between these models is that in the first one the spotlight only passively blocks the irrelevant information, while in the second model it actively directs the "processing efforts" to the chosen region.

Fig. 3(a) and Fig 3(b) visually clarify the difference between the spatial filtering and spotlight attention.

A conventional view of the spotlight model assumes that only a single region of interest is processed at a certain time point and supposes smooth movement to other regions of interest. Later versions of the spotlight model assume that the attentional spotlight can be divided between several regions in space. In addition, the latter support the theory that the spotlight moves discretely from one region to the other.

3.2 Object-based Model

As reviewed above, the spotlight metaphor is useful for understanding how attention is deployed across space. However, this metaphor has serious limitations. A detailed analysis of spotlight model drawbacks can be found in [3]. Object-based attention model suit to more practical experiments in physiology and is based on the assumption that attention is referred to discrete objects in the visual field. However being more practical, in contrast to the spotlight model, where one would predict that two nearby or overlapping objects are attended as a single object, in the object-based model this divided attention between objects results in less efficient processing than attending to a single object. It should be noted that spotlight and object-based attention theories are not contradictory but rather complementary. Nevertheless, in many cases the object-based theory explains many phenomena better than the spotlight model does.

The object-based model is more complicated for implementation, since it requires objects' recognition, while the spotlight model only requires identifying the regions of interest, where the attentional resources will be concentrated for further processing.

4. The Proposed Architecture for the CMOS Tracking Imager

In this Section two possible architectures of tracking CMOS imagers are discussed. While these architectures seem similar, the first one employing the spatial filtering model and the second one employing the spotlight attention model. The operation of both imagers will be described with the reference to an example of input scene in FOV, as shown in Fig. 4. In this scene three different types of regions can be observed: (a) two regions consisting of large salient targets (stars), (b) a number of regions consisting of small salient stars and (c) regions that don't consist of any targets of interest. The observer is usually interested in tracking the targets mentioned in group (a), but sometimes there is interest in targets both from groups (a) and (b). Moreover, sometimes the observer is interested in tracking the targets from group (b) and the targets from group (a) serve as salient distractors. Note, the term real time tracking relates to the ability to calculate the center of mass (COM) coordinates of the tracked target in real time.

Fig.4 An example of input scene in FOV

4.1 The Spatial Filtering Based Architecture

Fig. 5 shows the architecture of the CMOS tracking image system based on the spatial filtering model. The proposed sensor has two modes of operation: the fully autonomous mode and the semi-autonomous. In the fully autonomous mode all functions required for target tracking are performed by the sensor (at the pixel level and by on-chip image processing). The only data transmitted to the processor/DSP in this case is the tracked targets coordinates. This mode is very efficient by means of bottleneck problems solution; however it allows less flexibility and influence on the tracking process by the user. In the semi-autonomous mode, part of the functionality is performed in the circuit level and the chip level and part of processing is done by the processor/DSP. In this case feedback from the processor/DSP to the sensor exists and more flexibility is achieved; however more data flows from the sensors to the processor/DSP and back and more processor/DSP resources are used to complete the real time tracking (see Fig. 5).

The real time targets tracking is accomplished in two stages of operation: target acquisition and target tracking. In the acquisition mode N most salient targets of interest (the number of targets N can be predefined by the systems or can be user-defined) in the FOV are found. Then, N windows of interest with programmable size around the targets are defined, using the control logic block. These windows define the active regions, where the subsequent processing will occur, similar to the flexible spotlight size in the biological systems. In the tracking stage, the system sequentially attends only to the previously chosen regions, while completely inhibiting the dataflow from the other regions. This way the system based on the spatial model of attention allows distractors elimination, oppositely to a case of the spotlight model. According to the spotlight model appearance of the additional "salient" targets during the tracking of given targets of interest causes temporary or even permanent loss of the desired target.

Thanks to the control logic and to the CMOS imager flexibility, the proposed concept permits choosing the attended regions in the desired order, independent on the targets saliency. In addition it allows shifting the attention from one active region to the other, independent of the distance between the targets.

As can be seen more information is transmitted from the sensor array to the on-chip system-processing block during the acquisition mode than during the tracking mode. The reason for this is that during the acquisition mode the whole image is captured and transmitted to the on-chip image-processing block for further processing. In case of the fully autonomous mode of operation, the on-chip processing finds the center of mass coordinates of all targets of interest and windows of interest are defined by the control logic. In a case of semi-autonomous mode, some processing is performed in the on-chip processing block and the data is transmitted to the processor/DSP for further processing and windows of interest definition. Note, the acquisition mode is required only once at the beginning of tracking. During the tracking mode, only information from the chosen windows of interest is transmitted, dramatically reducing the bottleneck problem between the sensor and on-chip image-processing block and between the imager chip to processor/DSP.

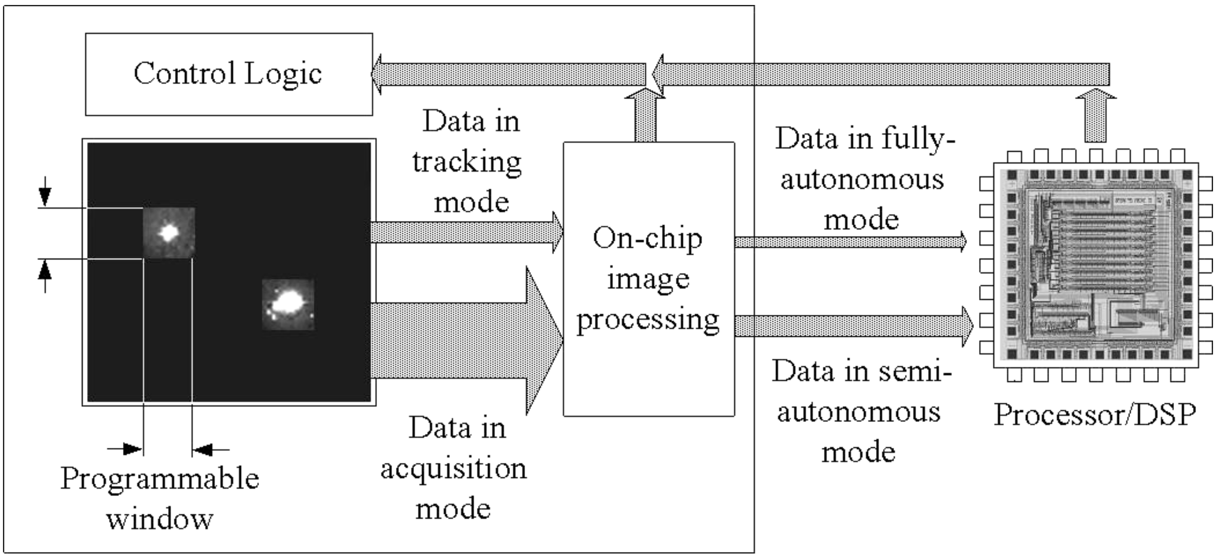

Fig.5 Architecture of the CMOS tracking image system based on the spatial filtering model

4.2 The Spotlight Based Architecture

Fig. 6 shows the architecture of the CMOS tracking image system based on the spotlight model. The principle of this system is very similar to the concept, presented in sub-section 4.1.

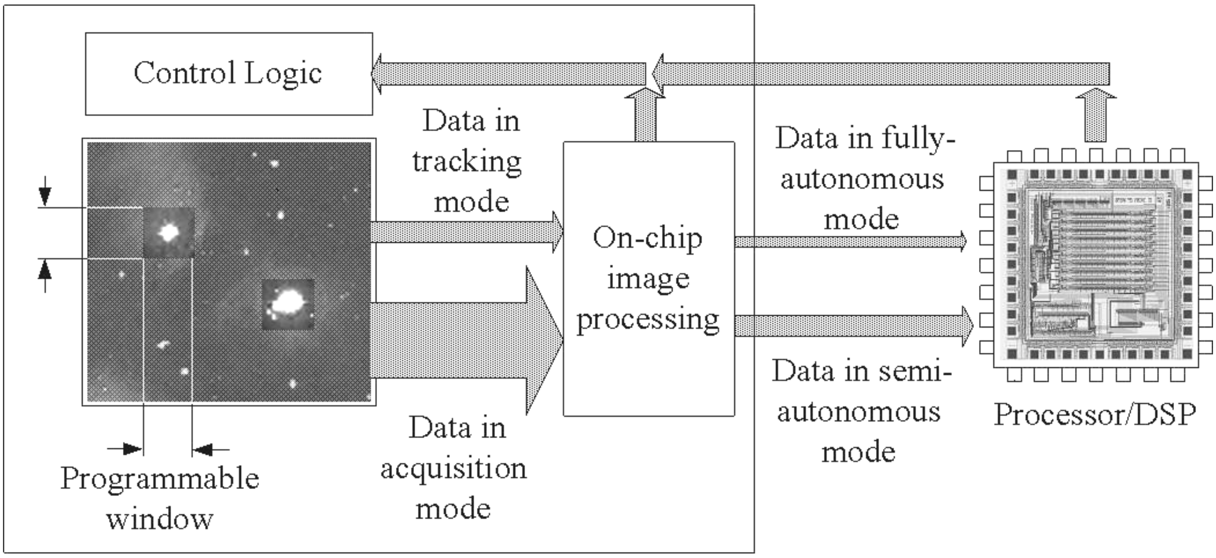

Fig.6 Architecture of the CMOS tracking image system based on the spotlight model

However, there is one important difference. While the spatial filter based imager filters all information that does not fall into the windows of interest, the spotlight based system, only reduces the amount of information transmitted from these regions for further processing. This is performed by a unique feature of CMOS image sensors that allow implementation of adaptive multiple resolution sensors [15]. As can be seen in Fig. 5, the two most important regions of interest are captured with full resolution (in the same way like in the system presented on sub-section 4.1), while all other regions are captured with reduced resolution. On one hand this allows to change the defined windows of interest according to the events in the FOV. On the other hand, this architecture is still significantly reduces the amount of information transmission from the sensor.

5. Conclusions

Two architectures of tracking image systems were proposed. Both the data transmission bottleneck and the performance bottleneck are reduced in the proposed imagers due to employing the spatial filtering and spotlight models of attention, found in biological systems. The proposed imagers can be easily implemented in a standard CMOS technology. Both imagers can operate in the full autonomous and semi-autonomous modes of operation. A brief description of the spatial and object-based models of attention was presented and an explanation of the proposed image systems operation was provided. Further research includes implementation of the proposed sensors in an advanced CMOS technology.

Acknowledgements

We would like to thank Alexander Spivakovsky and Evgeny Artyomov for their helpful suggestions during the preparation of this work. We also would like to thank the Israeli Ministry of Science and Technology for funding this project.

Bibliography

[1] O. Yadid-Pecht, R. Etienne-Cummings. CMOS imagers: from phototransduction to image processing. Kluwer Academic Publishers, 2004.
[2] R. William, H. Engineering. Using FPGAs for DSP Image Processing. FPGA journal, available http://www.fpgajournal.com/articles/imaging_hunt.htm
[3] Chun, M. M., & Wolfe, J. M. Visual Attention. In E. B. Goldstein (Ed.), Blackwell's Handbook of Perception, Vol. Ch 9, pp. 272-310). Oxford, UK: Blackwell, 2001.
[4] R. W. Remington, L. Pierce. Moving attention: Evidence for Time-invariant shifts of visual selection attention. Perception and Psychophysics, vol. 35, pp. 393-399, 1984.

[5] J. Moran, R. Desimone. Selective attention gates visual processing in the Extrastriate Cortex. Science, vol. 229, pp. 784-787, 1985.
[6] G. Sperling, E. Weichselgartner. Episodic theory of the dynamics of spatial attention. Psychological review, vol. 102, no. 3, pp. 503-532, 1995.
[7] A. Treisman. Features and targets: The fourteenth Barlett memorial lecture. Quarterly Journal of Experimental Psychology, 40A, pp. 201-237, 1988.
[8] A. Treisman. Feature binding, attention and target perception. Philosophical Transactions of the Royal Society of London, vol. 353, pp. 1295-1306, 1998.
[9] A. Treisman, G. Gelade. A feature-integration theory of attention. Cognitive Psychology, vol. 12, pp. 97-136, 1980.
[10] A. Treisman, D. Kahneman, J. Burkell. Perceptual targets and the cost of filtering. Perception & Psychophysics, vol. 33, pp. 527-532, 1983.
[11] S. Yantis. Targets, attention, and perceptual experience in "Visual Attention". R. D. Wright (Eds.), pp. 187-214. Oxford, NY: Oxford University Press, 1998.
[12] M. I. Posner. Orienting of attention. Quarterly Journal of Experimental Psychology, vol. 32,pp. 3-25, 1980.
[13] C. Koch, B. Mathur. Neuromorphic vision chips. IEEE Spectrum, vol. 33, pp. 38–46, May 1996
[14] C. Koch, S. Ullman. Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry. Human Neurobiology, vol. 4, pp. 219–227, 1985.
[15] E. Artyomov, O. Yadid-Pecht. Adaptive Multiple Resolution CMOS Active Pixel Sensor. ISCAS 2004, Vancouver, Canada, May 2004, Vol. IV, pp. 836-839.

Authors' Information

Alexander Fish – The VLSI Systems Center, Ben-Gurion University, Beer Sheva, Israel; e-mail: afish@ee.bgu.ac.il

Orly Yadid-Pecht – Dept of Electrical and Computer Engineering, University of Calgary, Alberta, Canada; e-mail: oyp@ee.bgu.ac.il

EXAMINATION OF PASSWORDS' INFLUATION ON THE COMPRESSING PROCESS OF NON-ENCRYPTED OBJECTS

Dimitrina Polimirova-Nickolova, Eugene Nickolov

Abstract: *The principal methods of compression and different types of non-encrypted objects are described. An analysis is made of the results obtained from examinations of the speed of compression for objects when using passwords with different length. The size of the new file obtained after compression is also analyzed. Some evaluations are made with regard to the methods and the objects used in the examinations. In conclusion some deductions are drawn as well as recommendations for future work.*

Keywords: *Password, Methods of Compression, Non-Encrypted Applications, Archive Programs, Level of Compression, File Extensions, Information Security.*

ACM Classification Keywords: *D.4.6 Security and Protection: information flow controls*

The Situation

The information technologies' progress leads to increasing need of creation and use of compressed objects. With regards to this, examination and analysis are made of different methods of compression and their varieties, which purpose is the creation of high-speed, effective compression of information flows working in real time.