JANUARY 2004
VOL. 14, NO. 1

# ELECTRONIC IMAGING

SPIE International Technical Group Newsletter

## Aliasing in digital cameras

Aliasing arises in all acquisition systems when the sampling frequency of the sensor is less than twice the maximum frequency of the signal to be acquired. We usually consider the spatial frequencies of the visual world to be unlimited, and rely on the optics of the camera to impose a cut-off. Thus, aliasing in any camera can be avoided by finding the appropriate match between the optics' modulation transfer function (MTF) and the sampling frequency of the sensor. In most digital cameras, however, the focal plane irradiance is additionally sampled by a Color Filter Array (CFA) placed in front of the sensor, composed of a mosaic of color filters. Consequently, each photo site on the sensor has only a single chromatic sensitivity. Using a CFA, a method first proposed by Bayer,[1] allows using one single sensor (CCD or CMOS) to sample color scenes. Missing colors are subsequently reconstructed, using a so-called demosaicing algorithm, to provide a regular three-color-per-pixel image.

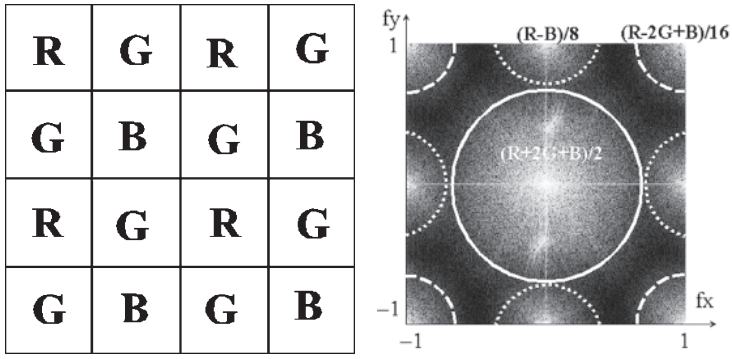



*Figure 1. (Left) An example of a color filter array, the Bayer Color Filter Array (Right) The Fourier representation of an image acquired with the Bayer Color Filter Array.*

In a CFA image acquisition system, a match between the optics' MTF and the sensor's sampling frequency is more difficult to establish because the sampling frequencies generally vary for each color (i.e. filter type). In the Bayer CFA, for example, there are twice as many green as red and blue filters (Figure 1a), resulting in different sampling frequencies for green and red/blue. Additionally, the horizontal and vertical sampling frequency for the green pixels is different from the diagonal frequency.

Thus, Greivenkamp[2] and Weldy[3] proposed an optical system called a birefringent lens that has varying spatial MTFs depending on wavelength. With such a lens it is possible to design a camera where the MTF of the optics matches the sampling frequency of each filter in the CFA. Thus, a color image could be reconstructed without artifacts. In practice, however, this method has not yet been applied because the resulting images are too blurry. They have a spatial resolution far lower than the resolving power of modern CCD or CMOS sensors. Most studies have therefore concentrated on how to reconstruct aliased images resulting from CFA camera systems where the optics is designed to pass high spatial frequencies.

If the captured scene has high spatial frequencies, the demosaiced image can contain visible artifacts. Depending on the scene content and the specific demosaicing algorithm used for reconstruction, they are more or less visible. In general, aliased sig-

*Continues on page 8.*

***Special Issue on:***

**Smart Image Acquisition and Processing**

*Guest Editors*

**François Berry,** Univ. Blaise Pascal

**Michael A. Kriss,** Sharp Lab. of America

### Newsletter now available on-line

Technical Group members are being offered the option of receiving the Electronic Imaging Newsletter electronically. An e-mail is being sent to all group members with advice of the web location for this issue, and asking members to choose between the electronic and printed version for future issues. If you are a member and have not yet received this message, then SPIE does not have your correct e-mail address.

To receive future issues electronically please send your e-mail address to:
`spie-membership@spie.org`
with the word `EI` in the subject line of the message and the words `electronic version` in the body of the message.

If you prefer to receive the newsletter in the printed format, but want to send your correct e-mail address for our database, include the words `print version preferred` in the body of your message.

# Smart camera and active vision: the active-detector formalism

Active vision techniques attempt to simulate the human visual system. In human vision, head motion, eye jerks and motions, and adaptation to lighting variations, are important to the perception process. Active vision, therefore, simulates this power of adaptation. Despite major shortcomings that limit the performance of vision systems—sensitivity to noise, low accuracy, lack of reactivity—the aim of active vision is to develop strategies for adaptively setting camera parameters (position, velocity, ...) to allow better perception. Pahlavan proposed that these parameters be split into four categories: optical parameters, for mapping the 3D world onto the 2D image surface; sensory parameters, for mapping from the 2D image to the sampled electrical signal; mechanical parameters, for the positioning and motion of the camera); and algorithmic integration to allow control of these parameters.

In the active approach to perception we assume that the outside world serves as its own model. Thus, perception involves exploring the environment allowing many traditional vision problems to be solved with low complexity algorithms. Based on this concept, the methodology used in this work involves integrating imager control in the perception loop and, more precisely, in early vision processes. This integration allows us to design a reactive-vision sensor. The goal is to adapt sensor attitude to environment evolution and the task currently being performed. Such smart cameras allow basic processing and the selection of relevant features close to the imager. This faculty reduces the significant problem of sensor communication flow. Further, as its name suggests, an active vision system actively extracts the information it needs to perform a task. By this definition, it is evident that one of its main goals is the selection of windows of interest (WOI) in the image and the concentration of processing resources on it: the notion of local study becomes predominant.

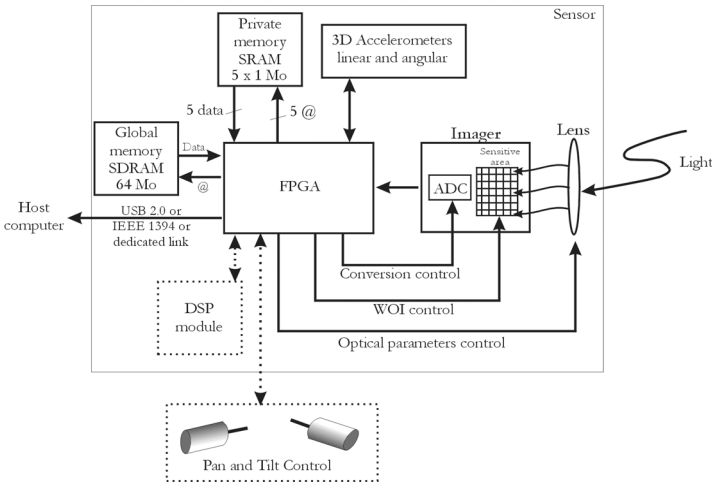



*Figure 1. Global architecture of the sensor. The blocks drawn with dashed lines represent optional modules that are not currently implemented.*

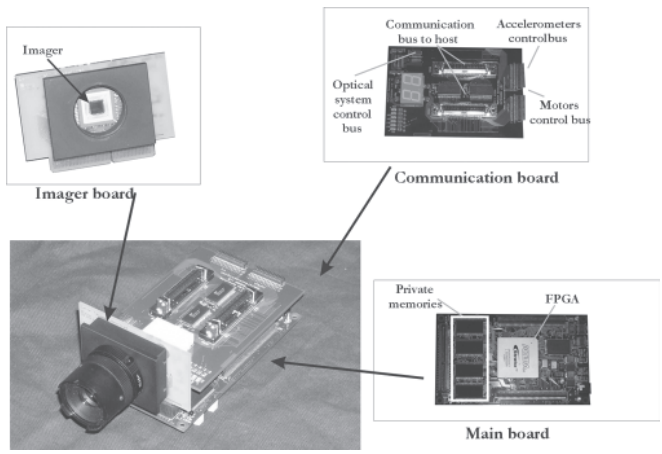



*Figure 2. Prototype of the sensor.*

Another important feature is the control of sensing parameters. As explained above, active vision devices generally focus on optical parameters (such as zoom, focus, and aperture), mechanical parameters (such as pan and tilt) and sometimes algorithmic integration (for example, early biologically-inspired vision systems). Thus, the visual sensor requires its own level of autonomy in order to: perform pre-processing, like the adjustment of sensing parameters or automatic WOI tracking; reduce the high-level system load; and/or reduce communication flow between the sensor and the host system.

### *Active detection*

Here, we include sensing parameters in the perception loop by introducing the notion of active detectors that control all levels of perception flow. These consist of hardware elements—such as a use of a sub-region of the imager or hardware implementation using a dedicated architecture—and software elements such as control algorithms.

An active detector can be viewed as a set of 'visual macro functions' where a visual task in decomposed into several subtasks. This is similar to Ullman's work[1] on the notion of a collection of 'visual routines' representing different kinds of basic image-processing sub functions. These can be used in goal-directed programs to perform elaborate tasks. By contrast, the active detector consists of both hardware and software. Thus, in this approach, the sensor has a key role in the perception process, its task more important than just performing image pre-processing. As a result, the hardware architecture and implementation are vital.

### *Smart architecture based on FPGA and CMOS imaging*

Our work is based on the use of a CMOS imager that allows full random-access readout and a massive FPGA architecture. There are several options for the choice of an imager and a processing unit. It is important to consider an active detector as a visual control loop: the measure is the image and the system to control is the sensor. For a given visual problem, the active detector must optimize and serve the sensor in order to achieve the task. For this reason, the architecture of this active vision sensor can be viewed as a set of parallel control loops where the bottleneck is the imager. Indeed, actual CMOS imagers have a sequential behavior and their acquisition rates are slowed in comparison with actual dedicated architectures performances.

The global architecture shown in Figure 1

*Continues on page 9.*

# Reflectance-sensitive retina

It is well established that intelligent systems start with good sensors. Attempts to overcome sensor deficiencies at the algorithmic level alone inevitably lead to inferior and unreliable overall-system performance: inadequate or missing information from the sensor cannot be made up in the algorithm. For example, in image sensors there is nothing we can do to recover the brightness of the image at a particular location once that pixel saturates. Sensors that implement some processing at the sensory level—computational sensors—may provide us with a level of adaptation that allows us to extract environmental information that would not be obtainable otherwise.

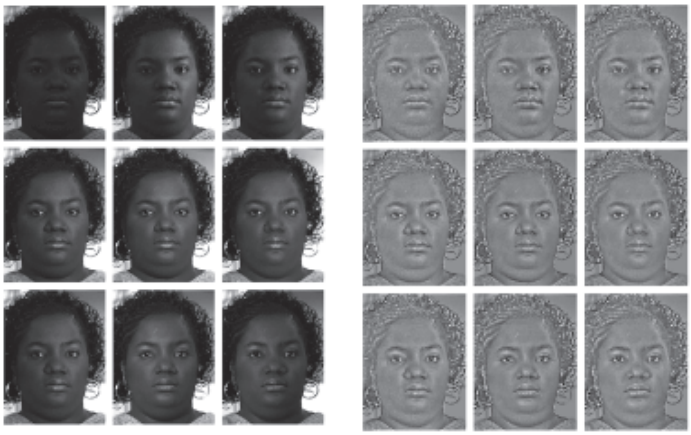
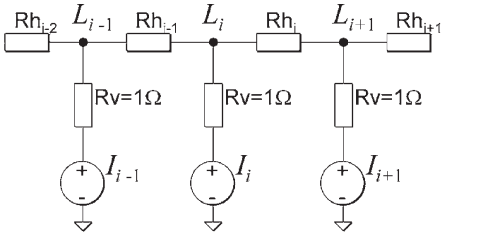


*Figure 1. The reflectance-sensitive retina removes illumination-induced variations (simulation results). The input images on the left are taken under varying illumination directions, resulting in substantial appearance changes. Our reflectance-recovery method largely removes these, resulting in virtually uniform appearance across different illumination conditions as shown on the right.*

Most present and future vision applications—including automotive, biometric, security, and mobile computing—involve unconstrained environments with unknown and widely-varying illumination conditions. Even when an image sensor is not saturated, the vision system has to account for object appearance caused by variations in illumination. To illustrate this, the left panel of Figure 1 shows a set of face images captured under varying illumination directions by a CCD camera. Even to a human observer, these faces do not readily appear to belong to the same person. We have recently introduced a reflectance perception model[1] that can be implemented at the sensory level and which largely removes the illumination variations as shown in the right panel of Figure 1. These images appear to be virtually identical. Using several standard face-recognition algorithms, we have shown that recognition rates are significantly improved when operating on the images whose variations due to the illumination are reduced by our method.[2]

In the most simplistic model, image intensity $I(x,y)$ is a product of object reflectance $R(x,y)$ and the illumination field $L(x,y)$, that is $I(x,y)=R(x,y)L(x,y)$. An illumination pattern $L$ is modulated by the scene reflectance $R$ and together they form the radiance map that is collected into a camera image $I$. $R$ describes the scene. In essence, $R$ is what we care about in computer vision. When the illumination field $L$ is uniform, $I$ is representative of $R$. But $L$ is rarely uniform. For example, the object may occlude light sources and create shadows.

Obviously, estimating $L(x,y)$ and $R(x,y)$ from $I(x,y)$ is an ill-posed problem. In many related illumination compensation methods, including Retinex,[1,4] a smooth version of the image $I$ is used as an estimate of the illumination field $L$.




$$J(I)=\sum_{i\in D}(I_i-L_i)^2+\lambda\frac{1}{R_{h_i}}|\nabla L_i|^2,\quad R_{h_i}=\frac{|\nabla I_i|}{I_i}$$

*Figure 2. The resistive grid that minimizes energy J(I), therefore finding the smooth version of the input image I. Perceptually important discontinuities are preserved because the horizontal resistors are controlled with the local Weber-Fechner contrast.*

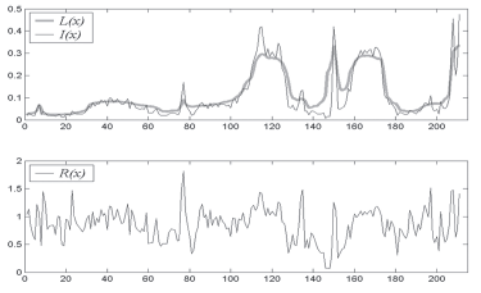



*Figure 3. Horizontal intensity line profiles through the middle of subject's eyes in top middle picture of Figure 1. The thin black line in the top graph is the original image'I(x,y), the thick gray line is the computed L(x,y), and the bottom graph is R(x,y) = I(x,y)/L(x,y).*

If this smooth version does not properly account for discontinuities, objectionable 'halo' artifacts are created in $R$ along the sharp edges in the image.

In our method, $L$ is estimated with the resistive network shown in Figure 2. Here, we use a one-dimensional example to keep the notation simple. The image pixel values are supplied as voltage sources and the solution for $L$ is read from the nodal voltages. To preserve discontinuities, the horizontal resistors are modulated proportionally to the Weber-Fechner contrast[5] between the two points interconnected by the horizontal resistor. Therefore, the discontinuities with large Weber-Fechner contrast will have a large resistance connecting the two points: smoothing less and allowing voltage at those two points to be kept further apart from each other. Formally, in the steady-state, the network minimizes the energy it dissipates as expressed by equation $J(I)$ shown in Figure 2. The first term is the energy dissipated on the vertical resistors $Rv$; the second term is the energy dissipated on the horizontal resistors $Rh$.

Once $L(x,y)$ is computed, the $I(x,y)$ is divided to produce $R(x,y)$. Figure 3 illustrates this process. It can be observed that the reflectance variations in shadows are amplified and 'pulled up' to the level of reflectance variations in the brightly-illuminated areas. All the details in the shadow region, which are not 'visible' in the original, are now clearly recognizable. We are currently designing an image sensor that implements this form of adaptation on the sensor chip before the signal is degraded by the readout and quantization process.

**Vladimir Brajovic**
The Robotics Institute
Carnegie Mellon University, USA
E-mail: brajovic@cs.cmu.edu

**References**

1. V. Brajovic, *A Model for Reflectance Perception in Vision,* **Proc. SPIE, Vol. 5119,** 2003, *to appear*.
2. R. Gross and V. Brajovic, *An Image Pre-processing Algorithm for Illumination Invariant Face Recognition,* **4th Int'l Conf. on Audio- and Video-Based Biometric Person Authentication,** 2003.
3. E. H. Land and J. J. McCann, *Lightness and Retinex theory,* **J.O.S.A 61** (1), pp. 1-11, January 1971.
4. D. J. Jobson, Z. Rahman, and G. A. Woodell, *A multi-scale Retinex for bridging the gap between color images and the human observation of scenes,* **IEEE Trans. Images Processing 6** (7), pp. 965-976, 1997.
5. B. A. Wandel, **Foundations of Vision,** Sinauer, Sunderland, MA, 1995.

# Color processing for digital cameras

Digital capture is becoming mainstream in photography, and a number of new color processing capabilities are being introduced. However, producing a pleasing image file of a natural scene is still a complex job. It is useful for the advanced photographer or photographic engineer to understand the overall digital camera color processing architecture in order to place new developments in context, evaluate and select processing components, and to make workflow choice decisions.

The steps in digital camera color processing are provided below.[1] Note that in some cases the order of steps may be changed, but the following order is reasonable. Also, it is assumed that proprietary algorithms are used to determine the adopted white[2] and the color-rendering transform.[3]

- Analog gains (to control sensitivity and white balance, if used)
- Analog dark-current subtraction (if used)
- A/D conversion
- Encoding nonlinearity (to take advantage of noise characteristics to reduce encoding bit-depth requirements for raw data storage, if used)

*This is the first raw-image data-storage opportunity*

- Linearize (undo any sensor and encoding nonlinearities; optionally clip to desired sensor range)
- Digital dark-current subtraction (if no analog dark-current subtraction)
- Optical flare subtraction (if performed)

*This is the last raw image data storage opportunity (before significant lossy processing)*

- Digital gains (to control sensitivity and white balance, if used)
- White clipping (clip all channels to same white level; needed to prevent cross-contamination of channels in matrixing)
- Demosaic (if needed)
- Matrix (to convert camera color channels to scene color channels)

*This is the scene-referred image data storage opportunity (standard scene referred color encodings include scRGB[4] and RIMM/ERIMM RGB)[5]*

- Apply color rendering transform (to take scene colors and map them to pleasing picture colors)
- Apply transform to standard output-referred encoding.

*This is the output-referred image-data storage opportunity (standard output-referred color encodings include sRGB,[6] sYCC,[7] ROMM RGB)[8]*

*Figure 1. Top left: raw image data with analog dark-current subtraction and gamma = 2.2 encoding nonlinearity, but no analog gains. Top right: "camera RGB" image data, after flare subtraction, white balancing and demosaicing (displayed using gamma = 2.2). Bottom left: scene-referred image data, after matrixing to scRGB color space (displayed using gamma = 2.2). Bottom right: sRGB image data, after color rendering and encoding transforms.*

### Workflow choices

The primary workflow choice is the image state for storage or exchange. The standard options are 'raw', 'scene-referred' and 'output-referred'.[9] The advantage of storing the image earlier in the processing chain is that decisions about subsequent processing steps can be changed without loss, and more advanced algorithms can be used in the future. The most lossy steps are white-clipping and color rendering. However, the white-clipping step is not needed if the camera to scene matrix does not need to be applied (i.e. the camera color channels can be encoded as scene color channels without matrixing), and carefully-designed color rendering transforms can minimize color rendering loss.

The disadvantages of exchanging image data using the raw or scene-referred image states are that many of the decisions affecting the appearance of the final picture have not yet been made. Control of the appearance is relinquished to the downstream processing. If a raw file is exchanged, the white-balance and color-rendering choices made after exchange can produce a variety of results, some of which may be quite different from the photographer's intent. Scene-referred image data has undergone white balancing (so the overall color cast of the image is communicated), but the color rendering step offers many opportunities for controlling the final image appearance. Output-referred exchange enables the photographer to communicate the desired final appearance in the image file, thereby ensuring more consistent output.

Current open-image exchange supports output-referred color encodings like sRGB. Generally it is recommended that output-referred images be exchanged for interoperability, although a raw or scene-referred image may be attached to allow other processing choices to be made in the future, or by other parties after exchange.

### Proprietary component choices

These include the methods for: determining the adopted white, flare subtraction, demosaicing (if needed), determining the matrix from camera color to scene color,[10] and determining the color-rendering transform. A camera's color reproduction quality will depend on each of these components. Generally, it is good to consider them independently, though one component may partially compensate for deficiencies in another. For example, if flare subtraction is omitted, a saturation boost in the camera-color-to-scene-color matrix or the color-rendering step may help, but the quality obtained will generally not be as good. Also, if the job of one step is deferred to another, image-data exchange in open systems is degraded: there may be no standard way to communicate that some operation was deferred.

### Optional proprietary step: scene relighting

Some scenes have variable lighting, or very high dynamic ranges due to light sources or cavities in the scene. Proprietary scene re-lighting algorithms attempt to digitally even out the scene illumination.[11] This can make scenes look more like they do to a human observer, because

*Continues on page 8.*

# Real-time image processing in a small, systolic, FPGA architecture

The need for high-performance computer architectures and systems is becoming critical in order to solve real image-processing applications. The implementation of such systems has become feasible with microprocessor technology development: however, conventional processors cannot always provide the computational power to fulfill real-time requirements due to their sequential nature, the large amount of data, and the heterogeneity of computations involved. Moreover, new trends in embedded system design restrict further the use of complex processors: large processing power, reduced physical size, and low power consumption are hard constraints to meet.[1] On the other hand, the inherent data- and instruction-level parallelism of most image-processing algorithms can be exploited to speed things up. This can be done through the development of special-purpose architectures on a chip based on parallel computation.[2]

Within this context, our research addresses the design and development of an FPGA hardware architecture for real-time window-based image processing. The wide interest in window-based or data-intensive processing is due to the fact that more complex algorithms can use low-level results as primitives to pursue higher-level goals. The addressed window-based image algorithms include generic image convolution, 2D filtering and feature extraction, gray-level image morphology, and template matching.

Our architecture consists of a configurable, 2D, systolic array of processing elements that provide throughputs of over tens of giga operations per second (GOPs). It employs a novel addressing scheme that significantly reduces the memory-access overhead and makes explicit the data parallelism at a low temporal storage cost.[3] A specialized processing element, called a configurable window processor (CWP), was designed to cover a broad range of window-based image algorithms. The functionality of the CWPs can be modified through configuration registers according to a given application.

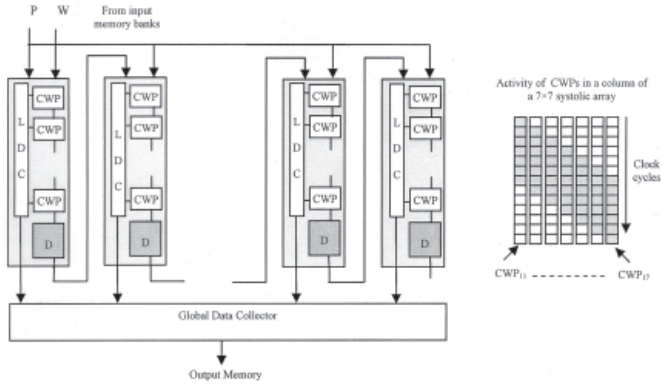



*Figure 1. Block diagram of the 2D systolic array of configurable window processors (CWPs) for window-based image processing. D is a delay line or shift register and LDC is a local data collector.*

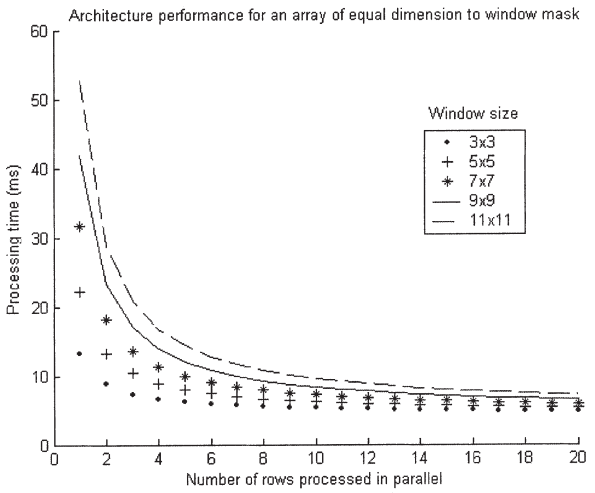



*Figure 2. Performance of the proposed architecture for a 512×512 gray level image with different window sizes.*

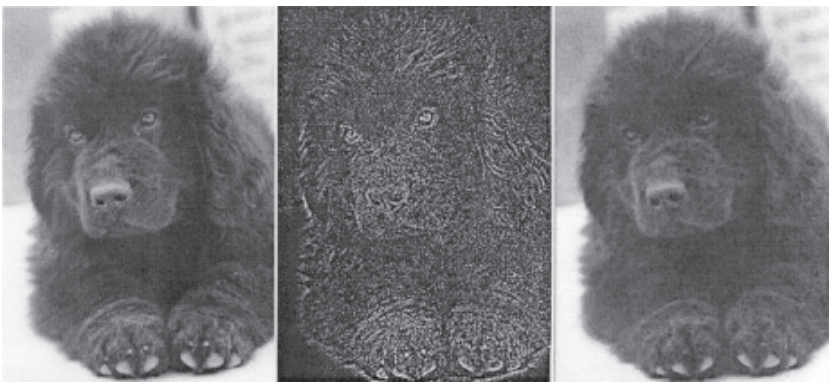


*Figure 3. Input image (left) and the output images for LoG filtering (middle) and gray-level erosion (right).*

Figure 1 shows a block diagram of the 2D systolic organization of the CWPs.[4] The systolic array exploits the 2D parallelism through the concurrent computation of window operations through rows and columns in the input image. For each column of the array there is a local data collector that collects the results of CWPs located in that column. The global data collector module collects the results produced in the array and sends them to the output memory.

As a whole, the architecture operation starts when a pixel from the input image is broadcast to all the CWPs in the array. Each concurrently keeps track of a particular window-based operation. At each clock cycle, a CWP receives a different window co-efficient *W*—stored in an internal register—and an input image pixel *P* that is common to all the CWPs. These values are used to carry out a computation, specified by a scalar function, and to produce a partial result of the window operation. The partial results are incrementally sent to the local reduction function implemented in the CWP to produce a single result when all the pixels of the window are processed. The CWPs in a column start working progressively: each a clock cycle delayed from the previous one as shown in Figure 1. The shadowed squares represent active CWPs in a given clock cycle.

A fully-parameterizable description of the modules of the proposed architecture was implemented using VHDL. The digital synthesis was targeted to a XCV2000E-6 VirtexE FPGA device. For an implemented 7×7 systolic-array prototype, the architecture provides a throughput of 3.16GOPs at a 60MHz clock frequency with a power con-

*Continues on page 9.*

# Cameras with inertial sensors

Inertial sensors attached to a camera can provide valuable data about camera pose and movement. In biological vision systems, inertial cues provided by the vestibular system play an important role, and are fused with vision at an early processing stage. Micromachining enables the development of low-cost single-chip inertial sensors that can be easily incorporated alongside the camera's imaging sensor, thus providing an artificial vestibular system. As in human vision, low-level image processing should take into account the ego motion of the observer. In this article we present some of the benefits of combining these two sensing modalities.

Figure 1 shows a stereo-camera pair with an inertial measurement unit (IMU), assembled with three capacitive accelerometers and three vibrating structure gyros. The 3D-structured world is observed by the visual sensor, and its pose and motion are directly measured by the inertial sensors. These motion parameters can also be inferred from the image flow and known scene features. Combining the two sensing modalities simplifies the 3D reconstruction of the observed world. The inertial sensors also provide important cues about the observed scene structure, such as vertical and horizontal references. In the system, inertial sensors measure resistance to a change in momentum, gyroscopes sense angular motion, and accelerometers change in linear motion. Inertial navigation systems obtain velocity and position by integration, and do not depend on any external references, except gravity.

The development of Micro-Electro-Mechanical Systems (MEMS) technology has enabled many new applications for inertial sensors beyond navigation, including aerospace and naval applications. Capacitive linear acceleration sensors rely on proof mass displacement and capacitive mismatch sensing. MEMS gyroscopes use a vibrating structure to measure the Coriolis effect induced by rotation, and can be surface micromachined providing lower-cost sensors with full signal-conditioning electronics. Although their performance is not suitable for full inertial navigation, under some working conditions or known system dynamics they can be quite useful.

In humans, the sense of motion is derived both from the vestibular system and retinal visual flow, which are integrated at very basic neural levels. The inertial information enhances the performance of the vision system in tasks such as gaze stabilisation, and visual cues aid spatial orientation and body equilibrium. There is also evidence that low-level human visual processing takes inertial cues into account, and that vertical and horizontal directions are important in scene interpretation. Currently-available MEMs inertial sensors have performances similar to the human vestibular system, suggesting their suitability for vision tasks.[1]

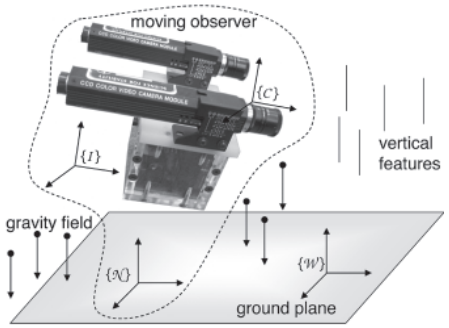



*Figure 1. Stereo cameras with an inertial measurement unit used in experimental work.*

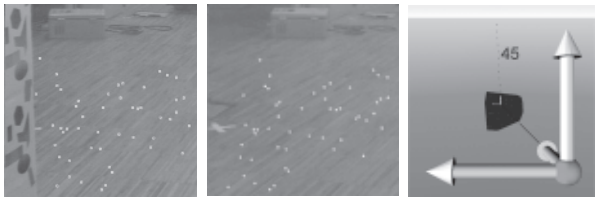



*Figure 2. Ground-plane 3D-reconstructed patch.*

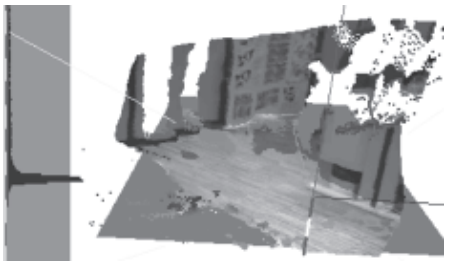


*Figure 3. Aligned depth map showing histogram for ground-plane detection.*

The inertial-sensed gravity vector provides a unique reference for image-sensed spatial directions. If the rotation between the inertial and camera frames of reference is known, the orthogonality between the vertical and the direction of a level plane image vanishing point can be used to estimate camera focal distance.[1] When the rotation between the IMU and camera is unknown from construction, calibration can be performed by having both sensors measuring the vertical direction.[2] Knowing the vertical-reference and stereo-camera parameters, the ground plane is fully determined. The collineation between image ground-plane points can be used to speed up ground-plane segmentation and 3D reconstruction (see Figure 2).[1] Using the inertial reference, vertical features starting from the ground plane can also be segmented and matched across the stereo pair, so that their 3D position is determined.[1]

The inertial vertical reference can also be used after applying standard stereo-vision techniques. Correlation-based depth maps obtained from stereo can be aligned and registered using the vertical-reference and dynamic-motion cues. In order to detect the ground plane, a histogram in height is performed on the vertically-aligned map, selecting the lowest local peak (see Figure 3). Taking the ground plane as a reference, the fusion of multiple maps reduces to a 2D translation and rotation problem. The dynamic inertial cues can be used as a first approximation for this transformation, providing a fast depth-map registration method.[3] In addition, inertial data can be integrated into optical flow techniques. It does this by compensating camera ego motion, improving interest-point selection, matching the interest points, and performing subsequent image-motion detection and tracking for depth-flow computation. The image focus of expansion (FOE) and centre of rotation (COR) are determined by camera motion and can both be easily found using inertial data alone, provided that the system has been calibrated. This information can be useful during vision-based navigation tasks.

Studies show that inertial cues play an important role in human vision, and that the notion of the vertical is important at the first stages of image processing. Computer-vision systems for robotic applications can benefit from low-cost MEMS inertial sensors, using both static and dynamic cues. Further studies in the field, as well as bio-inspired robotic applications, will enable a better understanding of the underlying principles. Possible applications go beyond robotics, and include of artificial vision and vestibular bio implants.

**Jorge Lobo and Jorge Dias**
Institute of Systems and Robotics
Electrical and Computer Engineering Department
University of Coimbra, Portugal
E-mail: {jlobo, jorge}@isr.uc.pt

**References**

1. J. Lobo and J. Dias, *Vision and Inertial Sensor Cooperation, Using Gravity as a Vertical Reference,* **IEEE Trans. on Pattern Analysis and Machine Intelligence 25** (12), 2003.
2. J. Alves, J. Lobo, and J. Dias, *Camera-Inertial Sensor modelling and alignment for Visual Navigation,* **Proc. 11th Int'l Conf. on Advanced Robotics,** pp. 1693-1698, 2003.
3. J. Lobo and J. Dias, *Inertial Sensed Ego-motion for 3D vision,* **Proc. 11th Int'l Conf. on Advanced Robotics,** pp. 1907-1914, 2003.

# Color artifact reduction in digital, still, color cameras

The tessellated structure of the color filter array overlaid on CMOS/CCD sensors in commercial digital color cameras requires the use of a considerable amount of processing to reconstruct a full-color image. Starting with a tessellated color array pattern—the Bayer Color Array is popularly used, and some other common filter-array tessellations are shown in Figure 1—we need to reconstruct a three-channel image. Clearly, missing data (colors not measured at each pixel) needs to be estimated. This is done via a process called demosaicing that introduces a host of color artifacts. Broadly, these can be split into two groups: so-called zipper and confetti artifacts.[1] The former occur at locations in the image where intensity changes are abrupt, and the latter when highly intense pixels are surrounded by dark pixels (usually a result of erroneous sensors). These artifacts may be reduced through a series of 'good' choices that range from the lens system to the choice of a very dense sensor (lots of photosites): used in conjunction with processing steps for correction.

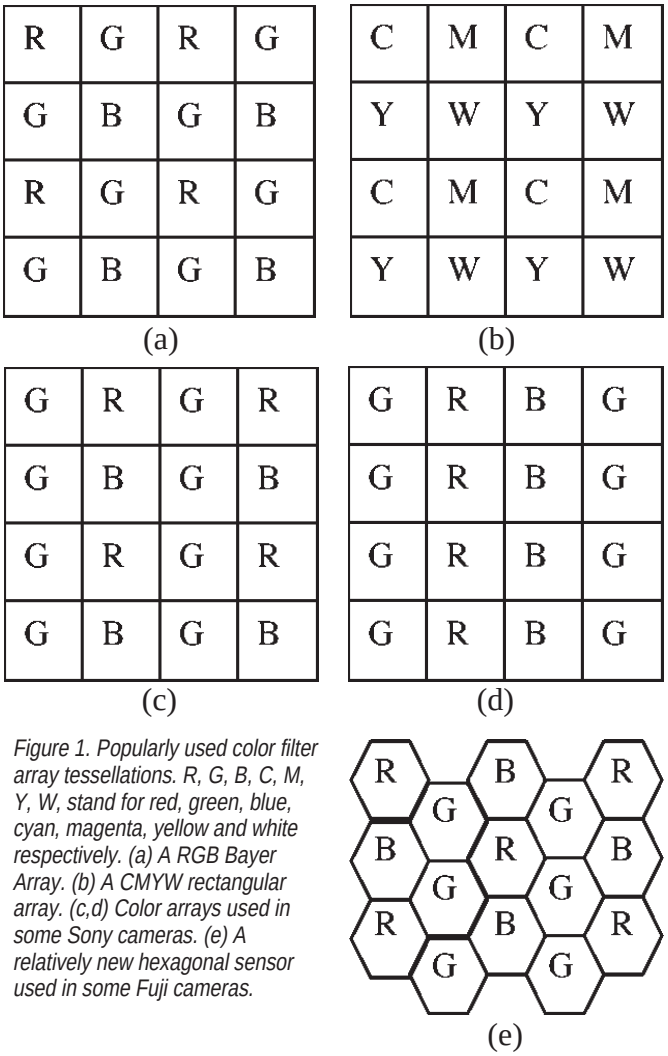



*Figure 1. Popularly used color filter array tessellations. R, G, B, C, M, Y, W, stand for red, green, blue, cyan, magenta, yellow and white respectively. (a) A RGB Bayer Array. (b) A CMYW rectangular array. (c,d) Color arrays used in some Sony cameras. (e) A relatively new hexagonal sensor used in some Fuji cameras.*

To remove these artifacts, the processing can either be done during or after the demosaicing step. Before we consider how these artifacts are removed/reduced, we need to bear in mind that the objective of commercial electronic photography is not so much the accurate reproduction of a scene, but a preferred or pleasing reproduction. In other words, even if there are errors introduced by the artifact removal stage, so long as the image 'looks' good, the consumer remains satisfied.[2]

As alluded to earlier, the reduction of artifacts could be performed during or after the demosaicing step. However, it is common to perform the artifact reduction at both stages of the image processing chain. Most demosaicing techniques make use of the fact that the human visual system is preferentially sensitive in the horizontal and vertical directions when compared to other directions (diagonal). When performing demosaicing, depending upon the strength of the intensity change in a neighborhood (horizontal, vertical, or diagonal) estimation kernels are used[3] that may be fixed or adaptive.[4,5] These are determined by operations over local neighborhoods—the goal being to interpolate along edges rather than across them (which leads to zipper errors).

Once a full-color image has been generated after demosaicing a filter-array image, the artifacts are either highly pronounced or relatively subdued depending on the technique used and image content. Most color-image-processing pipelines implement another collection of post-processing techniques to make the images appealing.

Most professional and high-end consumer cameras also have a post demosaicing noise-reduction step: usually a proprietary algorithm. However, a common algorithm used to reduce color artifacts is a median filter. Such artifacts usually have a salt-and-pepper type distribution over the image, for which the median filter is well suited. The human eye is known to be highly sensitive to sharp edges: we prefer sharp edges in a scene to blurred ones. Most camera manufacturers use an edge-enhancement step such as unsharp masking to make the image more appealing by reducing the low-frequency content in the image and enhancing the high frequency content. Another technique is called coring, used to remove detail information that has no significant contribution to image detail and behaves much like noise. The term coring originates from the manner in which the technique is implemented. Usually a representation of the data to be filtered is generated at various levels of detail, and noise reduction is achieved by thresholding (or 'coring') the transform coefficients computed at the various scales. How much coring needs to be performed (how high the threshold needs to be set) is a heuristic.

**Rajeev Ramanath and Wesley E. Snyder**
Department of Electrical and Computer Engineering
NC State University, USA
E-mail: rajeev.ramanath@ieee.org and wes@eos.ncsu.edu

**References**

1. R. Ramanath, W. E. Snyder, G. Bilbro, and W. A. Sander, *Demosaicking Methods for Bayer Color Arrays*, **J. of Electronic Imaging. 11** (3), pp. 306-315, 2002.
2. P. M. Hubel, J. Holm, G. D. Finlayson, and M. S. Drew, *Matrix calculations for digital photography*, **Proc. IS&T/SID 5th Color Imaging Conf.,** pp. 105-111, 1997.
3. J. E. Adams, *Design of practical Color Filter Array interpolation algorithms for digital cameras*, **Proc. SPIE 3028**, pp. 117-125, 1997.
4. R. Kimmel, *Demosaicing: Image reconstruction from color ccd samples*, **IEEE Trans. on Image Processing 8** (9), pp. 1221-1228, 1999.
5. R. Ramanath and W. E. Snyder, *Adaptive Demosaicking*, **J. of Electronic Imaging 12** (4), pp. 633-642, 2003.

## Aliasing in digital cameras

*Continued from the cover.*

nals cannot be recovered easily and only complicated methods using non-linear iterative processing[4] or prior knowledge[5] are able to effectively deal with this.

By studying the nature of aliasing in digital cameras, we have found a demosaicing solution that does not require excessive optical blur or complicated algorithms. As shown in Figure 1(b) and described formally elsewhere,[6] the Fourier spectrum of an image acquired with a Bayer CFA image has a particular pattern. Luminance (i.e. R + 2G + B) is localized in the center, and chrominance, composed of two opponent chromatic signals (R-G, -R+2G-B), are localized in the middle and corner of each side. The Fourier representation of a CFA image thus has the property of automatically separating luminance and opponent chromatic channels and projecting them in specific locations in the Fourier domain. Consequently, it is possible to directly estimate the luminance and chrominance signals with low- and high-pass filters, respectively, and then to reconstruct a color image by adding estimated luminance and estimated and interpolated chrominance.[6,7] Note, however, that luminance and opponent chromatic signals share the same two-dimensional Fourier space. Artifacts may result in the demosaiced image if their representations overlap (aliasing).

Using the Fourier representation thus also helps to illustrate the artifacts that may occur when applying any demosaicing algorithm: blurring occurs when luminance is estimated with a filter that is too narrow-band. False colors are generated when chrominance is estimated with a filter bandwidth that is too broad, resulting in high frequencies of luminance inside the chrominance signal. Grid effects occur when luminance is estimated with a bandwidth that is too broad, resulting in high frequencies of chrominance in the luminance signal. And, finally, water colors are generated when chrominance is estimated with a filter bandwidth that is too narrow. With many demosaicing algorithms, the two most visible effects are blurring and false color. For visual examples of the different artifacts, see Reference 7.

In general, algorithms that totally remove aliasing artifacts do not exist. However, in the case of a CFA image, the artifacts are not due to 'real' aliasing because they correspond to interference between luminance and chrominance: two different types of signals. This is certainly why many demosaicing methods work quite well. With our approach, one can optimally reconstruct the image without having recourse to any complicated de-aliasing methods. Our demosaicing-by-frequency-selection algorithms gives excellent results compared to other published algorithms and uses only a linear approach without any prior knowledge about the image content.[7] Also, our approach allows us to explicitly study demosaicing artifacts that could be removed by tuning spectral sensitivity functions,[8] optical blur, and estimation filters.

*Further information about this work and color illustrations are available at: http://ivrgwww.epfl.ch/index.php?name=EI_Newsletter*

**David Alleysson* and Sabine Süsstrunk†**
*Laboratory of Psychology and Neurocognition
Université Pierrre-Mendes, France
E-mail: David.Alleysson@upmf-grenoble.fr
†Audiovisual Communications Laboratory
School of Communications and Computing Sciences
Ecole Polytechnique Fédérale de Lausanne, Switzerland
E-mail: sabine.susstrunk@epfl.ch

**References**

1. B. E. Bayer, *Color Imaging Array,* **US Patent 3,971,065,** to Eastman Kodak Company, Patent and Trademark Office, Washington, D.C., 1976.
2. J. Greivenkamp, *Color dependant optical prefilter for the suppression of aliasing artifacts,* **Appl. Optics 29** (5), p. 676, 1990.
3. J. A. Weldy, *Optimized design for a single-sensor color electronic camera system,* **SPIE Optical Sensors and Electronic Photography 1071,** p. 300, 1989.
4. R. Kimmel, *Demosaicing: Image Reconstruction from Color Samples,* **IEEE Trans. Image Processing 8,** p. 1221, Sept. 1999.
5. D. H. Brainard, *Reconstructing Images from Trichromatic Samples, from Basic Research to Practical Applications,* **Proc. IS&T/SID 3rd Color Imaging Conf.,** p. 4, 1995.
6. D. Alleysson, S. Süsstrunk, and J. Hérault, *Color Demosaicing by estimating luminance and opponent chromatic signals in the Fourier domain,* **Proc. IS&T/SID 10th Color Imaging Conf.**, 2002.
7. http://ivrgwww.epfl.ch/index.php?name=EI_Newsletter
8. D. Alleysson, S. Süsstrunk, and J. Marguier, *Influence of spectral sensitivity functions on color demosaicing,* **Proc. IS&T/SID 11th Color Imaging Conf.**, 2003.

## Color processing for digital cameras

*Continued from page 4.*

the human visual system also attempts to compensate for scene lighting variability. Scene relighting algorithms should be evaluated based on how well they simulate real scene re-lighting, and the appearance of scenes as viewed by human observers. It is important to remember that re-lit scenes will then be color rendered; sometimes these two proprietary steps are combined to ensure optimal performance.

**Jack Holm**
Hewlett-Packard Company, USA
E-mail: jack.holm@hp.com

**References**

1. J. Holm, I, Tastl, L. Hanlon, and P. Hubel, *Color processing for digital photography,* **Colour Engineering: Achieving Device Independent Colour,** Phil Green and Lindsay MacDonald, editors, Wiley, 2002
2. P. M. Hubel, J. Holm, and G. D. Finlayson, *Illuminant Estimation and Color Correction,* **Colour Imaging in Multimedia,** Lindsay MacDonald, editor, Wiley, 1999.
3. J. Holm, *A Strategy for Pictorial Digital Image Processing (PDIP),* **Proc. IS&T/SID 4th Color Imaging Conf.: Color Science, Systems, and Applications,** pp. 194-201, 1996.
4. **IEC/ISO 61966-2-2:2003,** *Multimedia systems and equipment—Colour measurement and management—Extended RGB colour space—scRGB.*
5. **ANSI/I3A IT10.7466:2002,** *Photography—Electronic Still Picture Imaging—Reference Input Medium Metric RGB Color Encoding: RIMM RGB.*
6. **IEC 61966-2-1:1999,** *Multimedia systems and equipment—Colour measurement and management—Default RGB colour space—sRGB.*
7. **IEC 61966-2-1 Amendment 1: 2003.**
8. **ANSI/I3A IT10.7666:2002,** *Photography—Electronic Still Picture Imaging—Reference Output Medium Metric RGB Color Encoding: ROMM RGB.*
9. **ISO 22028-1:2003,** *Photography and graphic technology—Extended colour encodings for digital image storage, manipulation and interchange—Part 1: architecture and requirements.*
10. J. Holm, I. Tastl, and S. Hordley, *Evaluation of DSC (Digital Still Camera) Scene Analysis Error Metrics—Part 1,* **Proc. IS&T/SID 8th Color Imaging Conf.: Color Science, Systems, and Applications,** pp. 279-287, 2000.
11. J. McCann, *Lessons learned from mondrians applied to real images and color gamuts,* **Proc. IS&T/SID 7th Color Imaging Conf.: Color Science, Systems, and Applications,** pp. 1-8, 1999.

## Smart camera and active vision: the active-detector formalism

*Continued from page 2.*

presents the different modules of our device. The blocks drawn with dashed lines represent optional modules that are not currently implemented. The first prototype (see Figure 2) is composed of three parts: the imager, the main board, and the communications board. The first board includes a CMOS imager, analog/digital converter and optics. The main board is the core of the system: it consists of a Stratix from Altera; several private memory blocks; and, on the lower face of the board, an optional DSP module that can be connected for dedicated processing and an SDRAM module socket that allows the memory to be extended to 64 Mb. The communications board is connected to the main board and manages all communications with the host computer. On this card, we can connect a 3D accelerometer, zoom controller, and motor controller for an optional turret.

Our initial results show high speed tracking of a gray-level template (see Figure 2). According to the size of the window, the acquisition rate varies from 200-5500 frames per second.[2]

**François Berry and Pierre Chalimbaud**
LASMEA Laboratory, Université Blaise Pascal, France
**E-mail:** {berry, chalimba}@lasmea.univ-bpclermont.fr

**References**

1. S. Ullman, *Visual Routines,* **Cognition. 18,** p. 97-159, 1984.
2. P. Chalimbaud, F. Berry, and P. Martinet, *The task 'template tracking' in a sensor dedicated to active vision,* **IEEE Int'l Workshop on Computer Architectures for Machine Perception,** 2003.

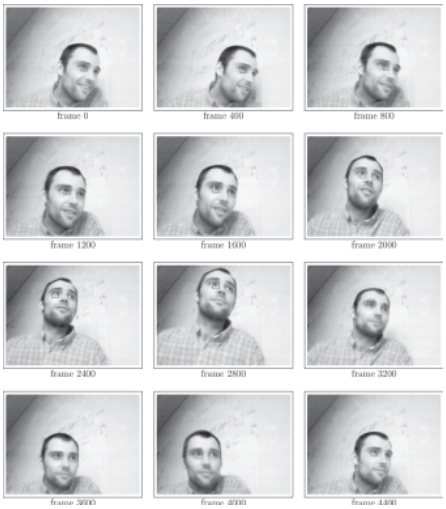



*Figure 3. High speed tracking of a gray-level template (32×32 at ~2000 frames per second).*

## Real-time image processing in a small systolic FPGA architecture

*Continued from page 5.*

sumption estimation of 1.56W. The architecture uses 6118 slices, i.e. around 30% of the FPGA. The architecture was validated on a RC1000-PP FPGA AlphaData board. The performance improvement on the software implementation running on a Pentium IV processor is more than an order of magnitude.

The processing times for a window-based operation on 512×512 gray-level images for different window sizes are plotted in Figure 2. The array was configured to use the same number of CWPs as the window size. For all the cases it was possible to achieve real-time performance with three to four rows processed in parallel. The processing time for a generic window-based operator with a 7×7 window mask on 512×512 gray-level input images is 8.35ms, thus the architecture is able to process about 120 512×512 gray-level images per second. Among the window-based image algorithms already mapped into and tested are generic convolution, gray-level image morphology and template matching. Figure 3 shows a test image and two output images for LoG filtering and gray-level erosion.

According to theoretical and experimental results, the architecture compares favorably with other dedicated architectures in terms of performance and hardware resource. Due to its configurable, modular, and scalable design, the architecture constitutes a platform to explore more complex algorithms such as motion estimation and stereo disparity computation, among others. The proposed architecture is well suited to be the computational core of a completely self-contained vision system due to its efficiency and compactness. The architecture can be coupled with a digital image sensor and memory banks on a chip to build a compact smart sensor for mobile applications.

**César Torres-Huitzil and Miguel Arias-Estrada**
Computer Science Department
INAOE, México
E-mail: {ctorres, ariasm}@inaoep.mx

**References**

1. J. Silc, T. Ungerer, and B. Robic, *A Survey of New Research Directions in Microprocessors,* **Microprocessor and Microsystems 34,** pp. 175-190, 2000.
2. N. Ranganathan, *VLSI & Parallel Computing for Pattern Recognition & Artificial Intelligence,* **Series in Machine Perception and Artificial Intelligence 18,** World Scientific Publishing, 1995.
3. Miguel Arias-Estrada and César Torres-Huitzil, *Real-time Field Programmable Gate Array Architecture for Computer Vision,* **J. Electronic Imaging 10** (1), pp. 289-296, January 2001.
4. César Torres-Huitzil and Miguel Arias-Estrada, *Configurable Hardware Architecture for Real-time Window-based Image Processing,* **Proc. FPL 2003,** pp. 1008-1011, 2003.

## Tell us about your news, ideas, and events!

If you're interested in sending in an article for the newsletter, have ideas for future issues, or would like to publicize an event that is coming up, we'd like to hear from you. Contact our technical editor, Sunny Bains (sunny@spie.org) to let her know what you have in mind and she'll work with you to get something ready for publication.

**Deadline for the next edition, 14.2, is:**

***19 January 2004:*** Suggestions for special issues and guest editors.

***26 January 2004:*** Ideas for articles you'd like to write (or read).

***26 March 2004:*** Calendar items for the twelve months starting June 2004.

# Calendar

## 2004

**IS&T/SPIE 16th International Symposium Electronic Imaging: Science and Technology**
***18–22 January***
*San Jose, California USA*
Program • Advance Registration Ends 17 December 2003
Exhibition
**http://electronicimaging.org/program/04/**

**Photonics West 2004**
***24–29 January***
*San Jose, California USA*
Featuring SPIE International Symposia:
- **Biomedical Optics**
- **Integrated Optoelectronic Devices**
- **Lasers and Applications in Science and Engineering**
- **Micromachining and Microfabrication**

Program • Advance Registration Ends 7 January 2004
Exhibition
**http://spie.org/Conferences/Programs/04/pw/**

**SPIE International Symposium Medical Imaging 2004**
***14–19 February***
*San Diego, California USA*
Program • Advance Registration Ends 5 February 2004
Exhibition
**http://spie.org/conferences/programs/04/mi/**

**SPIE International Symposium Optical Science and Technology**
**SPIE's 49th Annual Meeting**
***2–6 August***
*Denver, Colorado USA*
Call for Papers • Abstracts Due 5 January 2004
Exhibition
**http://spie.org/conferences/calls/04/am/**

**SPIE International Symposium ITCom 2004**
**Information Technologies and Communications**
***12–16 September***
*Anaheim, California USA*
Co-located with NFOEC

**26th International Congress on High Speed Photography and Photonics**
***20–24 September***
*Alexandria, Virginia USA*
Call for Papers • Abstracts Due 15 March 2004
**http://spie.org/conferences/calls/04/hs/**

**NIH Workshop on Diagnostic Optical Imaging and Spectroscopy**
***20–22 September***
*Washington, D.C. USA*
Organized by NIH, managed by SPIE

**SPIE International Symposium Photonics Asia 2004**
***8–11 November***
*Beijing, China*
Call for Papers
**http://spie.org/conferences/calls/04/pa/**

**For More Information Contact**
**SPIE • PO Box 10, Bellingham, WA 98227-0010 • Tel: +1 360 676 3290 • Fax: +1 360 647 1445**
**E-mail: spie@spie.org • Web: www.spie.org**

## *Space-variant image processing: taking advantage of data reduction and polar coordinates.*

*Continued from page 12.*

acquisition of frames per second is accelerated since the images are very small. The frame-grabber size is also dramatically reduced. Combined, these two effects make the exploitation of differential algorithms especially interesting. Such image-processing algorithms systematically apply simple operations to the whole image, computing spatial and temporal differences. These can be computationally intensive for large images and the simultaneous storage of several frames for computing temporal differences can be a hardware challenge. Log-polar image-data reduction can therefore contribute to the effective use of differential algorithms in real applications.[7]

In addition to the selective reduction of information, another interesting advantage of log-polar representation is related to polar coordinates. In this case, approaching movement along the optical axis in the sensor plane has only a radial coordinate. This type of movement is often present with a camera on top of a mobile platform like an autonomous robot. If the machine is moving along its optical axis, the image displacement due to its own movement has only a radial component. Thus, complex image-processing algorithms are simplified and accelerated.[3,7,8] Further, the hardware reduction achieved in storing and processing images, combined with the density of programmable devices, make possible a full image-processing system on a single chip.[9] This approach is especially well suited to systems with power consumption and hardware constraints. We would argue it is the natural evolution of the reconfigurable architectures employed for autonomous robotic navigation[7] systems.

*This work is supported by the Generalitat Valenciana under project CTIDIA/2002/142.*

**Jose A. Boluda and Fernando Pardo**
Departament d'Informàtica
Universitat de València, Spain
E-mail: Jose.A.Boluda@uv.es
**http://www.uv.es/~jboluda/**

**References**

1. M. Tistarelli and G. Sandini, *Dynamic aspects in active vision*. **CVGIP: Image Understanding 56** (1), p 108, 1992.
2. R. M. Hodgson and J. C. Wilson, *Log polar mapping applied to pattern representation and recognition,* **Computer Vision and Image Processing,** Ed. Shapiro & Rosenfeld, Academic Press, p. 245, 1992.
3. M. Tistarelli and G. Sandini, *On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow,* **IEEE Trans. on PAMI 15** (4), p. 401, April 1993.
4. R. S. Wallace, P. W. Ong, B. B. Bederson, and E. L. Schwartz, *Space-variant image processing,* **Intl. J. of Computer Vision 13** (1), p. 71, 1994.
5. F. Jurie, *A new log-polar mapping for space variant imaging: Application to face detection and tracking,* **Pattern Recognition 32** (5), p. 865, May 1999.
6. F. Pardo, B. Dierickx, and D. Scheffer, *Space-Variant Non-Orthogonal Structure CMOS Image Sensor Design,* **IEEE J. of Solid-State Circuits 33** (6), p. 842, June 1998.
7. J. A. Boluda and J. J. Domingo, *On the advantages of combining differential algorithms, pipelined architectures and log-polar vision for detection of self-motion from a mobile robot,* **Robotics and Autonomous Systems 37 (**4), p. 283, December 2001.
8. J. A. Boluda and F. Pardo, *A reconfigurable architecture for autonomous visual navigation,* **Machine Vision and Applications 13** (5-6), p. 322, 2003.
9. J. A. Boluda and F. Pardo. *Synthesizing on a reconfigurable chip an autonomous robot image processing system,* **Field Programmable Logic and Applications**, Springer Lecture Notes in Computer Science 2778, pp. 458-467, 2003.

# Join the SPIE/IS&T Technical Group

## ...and receive this newsletter

This newsletter is produced twice yearly and is available only as a benefit of membership of the SPIE/IS&T Electronic Imaging Technical Group.

IS&T—The Society for Imaging Science and Technology has joined with SPIE to form a technical group structure that provides a worldwide communication network and is advantageous to both memberships.

Join the Electronic Imaging Technical Group for US$30. Technical Group members receive these benefits:

- *Electronic Imaging* Newsletter
- SPIE's monthly publication, ***oe****magazine*
- annual list of Electronic Imaging Technical Group members

People who are already members of IS&T or SPIE are invited to join the Electronic Imaging Technical Group for the reduced member fee of US$15.

***Please Print*** ☐ Prof. ☐ Dr. ☐ Mr. ☐ Miss ☐ Mrs. ☐ Ms.

First (Given) Name _____ Middle Initial _____

Last (Family) Name _____

Position _____

Business Affiliation _____

Dept./Bldg./Mail Stop/etc. _____

Street Address or P.O. Box _____

City _____ State or Province _____

Zip or Postal Code _____ Country _____

Phone _____ Fax _____

E-mail _____

**Technical Group Membership fee is $30/year, or $15/year for full SPIE and IS&T Members.**

**Amount enclosed for Technical Group membership** **$** _____

❑ **I also want to subscribe to IS&T/SPIE's *Journal of Electronic Imaging* (JEI)** **$** _____
*(see prices below)*

**Total $** _____

❑ **Check enclosed.** Payment in U.S. dollars (by draft on a U.S. bank, or international money order) is required. Do not send currency. Transfers from banks must include a copy of the transfer order.

❑ **Charge to my:** ❑ **VISA** ❑ **MasterCard** ❑ **American Express** ❑ **Diners Club** ❑ **Discover**

Account # _____ Expiration date _____

Signature _____
*(required for credit card orders)*

**Reference Code: 3537**

| **JEI 2003 subscription rates (4 issues):** | **U.S.** | **Non-U.S.** |
|---|---|---|
| Individual SPIE or IS&T member* | $ 55 | $ 55 |
| Individual nonmember and institutions | $255 | $275 |

Your subscription begins with the first issue of the year. Subscriptions are entered on a calendar-year basis. Orders received after 1 September 2003 will begin January 2004 unless a 2003 subscription is specified.

***One journal included with SPIE/IS&T membership. Price is for additional journals.***

**Send this form (or photocopy) to:**
**SPIE • P.O. Box 10**
**Bellingham, WA 98227-0010 USA**
**Tel: +1 360 676 3290**
**Fax: +1 360 647 1445**
**E-mail: membership@spie.org**

**Please send me:**
☐ **Information about full SPIE membership**
☐ **Information about full IS&T membership**
☐ **Information about other SPIE technical groups**
☐ **FREE technical publications catalog**


# Electronic Imaging

The *Electronic Imaging* newsletter is published by SPIE—The International Society for Optical Engineering, and IS&T—The Society for Imaging Science and Technology. The newsletter is the official publication of the International Technical Group on Electronic Imaging.

| | |
|---|---|
| *Technical Group Chair* | Arthur Weeks |
| *Technical Group Cochair* | Gabriel Marcu |
| *Technical Editor* | Sunny Bains |
| *Managing Editor/Graphics* | Linda DeLano |

Articles in this newsletter do not necessarily constitute endorsement or the opinions of the editors, SPIE, or IS&T. Advertising and copy are subject to acceptance by the editors.

SPIE is an international technical society dedicated to advancing engineering, scientific, and commercial applications of optical, photonic, imaging, electronic, and optoelectronic technologies.

IS&T is an international nonprofit society whose goal is to keep members aware of the latest scientific and technological developments in the fields of imaging through conferences, journals and other publications.

**SPIE—The International Society for Optical Engineering,** P.O. Box 10, Bellingham, WA 98227-0010 USA. Tel: +1 360 676 3290. Fax: +1 360 647 1445. E-mail: spie@spie.org.

**IS&T—The Society for Imaging Science and Technology,** 7003 Kilworth Lane, Springfield, VA 22151 USA. Tel: +1 703 642 9090. Fax: +1 703 642 9094.




**EI**ONLINE

## Electronic Imaging Web Discussion Forum

You are invited to participate in SPIE's online discussion forum on Electronic Imaging. To post a message, log in to create a user account. For options see "subscribe to this forum."

You'll find our forums well designed and easy to use, with many helpful features such as automated email notifications, easy-to-follow 'threads,' and searchability. There is a full FAQ for more details on how to use the forums.

Main link to the new Electronic Imaging forum:

`http://spie.org/app/forums/tech/`

Related questions or suggestions can be sent to `forums@spie.org`.

The International Society for Optical Engineering

# Space-variant image processing: taking advantage of data reduction and polar coordinates.

The human retina exhibits a non-uniform photo-receptor distribution: more resolution at the center of the image and less at the periphery. This space-variant vision emerges as an interesting image acquisition system, since there is a selective reduction of information. Moreover, the log-polar mapping—as a particular case of space-variant vision—shows interesting mathematical properties that can simplify several widely-studied image-processing algorithms.[1-4] For instance, rotations around the sensor center are converted to simple translations along the angular coordinate, and homotheties (linear transformations) with respect to the center in the sensor plane become translations along the radial coordinate.

The sensor (with the space-variant density of pixels) and computational planes are called the retinal and cortical planes, respectively. The resolution of a log-polar image is usually expressed in terms of rings and number of cells (sectors) per ring. A common problem with this transformation is how to solve the central singularity: if the log-polar equations are strictly followed, the center would contain an infinite density of pixels that cannot be achieved. This problem of the fovea (the central area with maximum resolution) can be addressed in different ways: the central blind spot model, Jurie's model,[5] and other approaches that give special transformation equations for this central area. Figure 1 shows an example of a log-polar transformation. At the left there is a Cartesian image of 440×440 pixels; at the center is the same image after a log-polar transformation with a central blind spot that gives a resolution of 56 rings with 128 cells per ring. Notice there is enough resolution at the center to perceive the cat in detail. The rest of the image is clearly worse than the Cartesian version, but this is the periphery of the image. This retinal image occupies less than 8 kB: the equivalent Cartesian image is around 189 kB (24 times larger). The computational plane of the image is shown in Figure 1 (right).

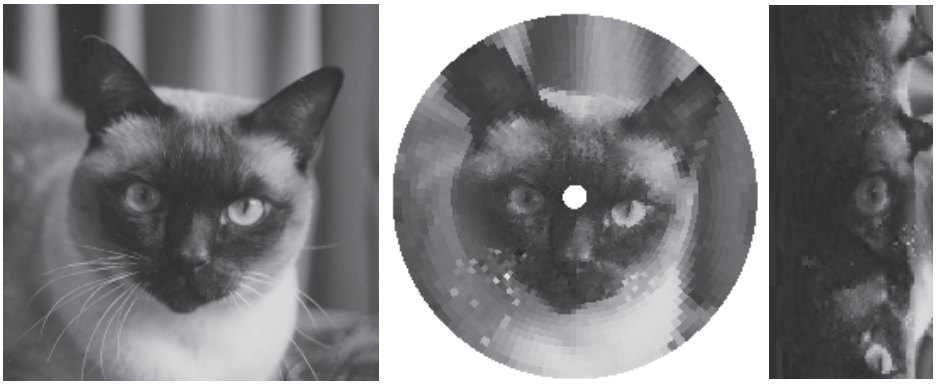


*Figure 1. Left: A 440×440 Cartesian image. Center: A 128×56 log-polar image. Right: The computational image.*

The best way to obtaining log-polar images depends on the available hardware and software. The simplest approach is to use software to transform a typical Cartesian image from a standard camera. This is done using the transformation equations between the retinal plane and the Cartesian plane. Since the transformation parameters can be tuned online, this solution is flexible. However, it can be an excessively-time-consuming effort if the computer must first process these images in order to perform another task. The other option is the purely-hardware solution: the log-polar transformation made directly from a sensor with this particular pixel distribution. An example of a log-polar sensor is a CMOS visual sensor designed with a resolution of 76 rings and 128 cells per ring.[6] The fovea is comprised of the inner 20 rings that follow a linear- (not log-) polar transformation to avoid the center singularity. This method fixes the image transformation parameters and is not flexible.

As an intermediate approach, a circuit that performs a Cartesian to log-polar image transformation can be implemented on a programmable device. This solution gives the advantage of speed while retaining flexibility: the transformation parameters can be changed on the fly. Moreover, the complexity and density of current reconfigurable devices represent a new trend in computer architecture, since it is possible to include microprocessors, DSP cores, custom hardware, and small memory blocks in a single chip.

The log-polar image data reduction has several positive consequences for the processing system. The first and most obvious is that the

*Continues on page 10.*

SPIE Society of Photo-Optical Instrumentation Engineers
P.O. Box 10 • Bellingham, WA 98227-0010 USA

Change Service Requested

**DATED MATERIAL**

Non-Profit Org.
U.S. Postage Paid
Society of
Photo-Optical
Instrumentation
Engineers