Digital Zoom Camera with Image Sharpening and Noise Suppression

Chaminda Weerasinghe, Magnus Nilsson, Serge Lichman and Igor Kharitonenko

Abstract —*The intention of zooming in a video camera is to obtain more detailed visual information than the full frame view. The optical zooming mechanisms provide this additional information, however, when optically zoomed, the camera no longer detects the full frame view, which was detected before zooming. Digital zooming can detect the full view at all times, however, fails to provide additional visual information when zoomed. In applications such as visual surveillance and manufacturing quality measurement, it is required to detect the full frame view at all times, however, the camera should be capable of providing additional visual information when zoomed to the region of interest (ROI). This paper presents a video camera system incorporating an autonomous pan, tilt and zoom functionality based on ROI decided upon motion detection capability, which provides more visual information in zoomed modes while detecting the full view at all times. The intelligent ROI decision can be made not only based on motion but also shape, color, sizes etc. of targeted objects for tracking. Image sharpening and noise reduction methods are also provided within the camera system. Standard testing results indicate that this camera system can resolve up to 800 TV lines and produces comparable color quality to other CCD based CCTV surveillance cameras, especially at good lighting conditions.*

Index Terms — Digital zoom, image sharpening, motion based pan tilt zoom and noise reduction.

I. INTRODUCTION

AUTOMATIC pan tilt zoom (PTZ) capability is a highly desired feature for cameras targeting surveillance and video monitoring applications. One of the popular solutions is to use a computer controlled PTZ moving camera [1]. However, this solution incurs high maintenance cost due to movable parts (servo motors) incorporated on the camera platform for mechanical pan and tilt capability. Optical zoom lenses are also expensive and difficult to maintain. Apart from the cost, region of awareness [2] is also an important factor in a camera deployed for surveillance and monitoring. Region of awareness (ROA) represents the field of view that is constantly monitored by the camera. At a particular instant in time, a camera with an optical zoom would have a small region of awareness, especially when fully zoomed to the region of interest (ROI). ROI is the area, which is displayed on the output video. Usually, ROA is expected to be much larger area compared to ROI. However, with an optical zoom camera, ROA=ROI at any point in time. In order to monitor a large ROA while zooming to a smaller ROI, electronic PTZ should be implemented. In this case, the ROA remains constant, while ROI is scaled to fit the output video frame size. ROI scaling up usually causes video blur and deterioration of image quality, compared to the optical zoom solution, which produces the same resolution and image quality at all zoom levels. One method of enhancing the electronic zoom resolution is to use a very large image sensor array [3] or deploy several cameras to capture the ROA [2]. Using several cameras incur additional cost and computational overhead in image stitching and registration as well as in camera calibration. Therefore, this paper presents a solution that uses a high-resolution image sensor together with a wide-angle lens to maximize the ROA.

Electronic zoom is capable of providing an ROI as large as the ROA and as small as that of the highest zoom level. Usually, when full ROA is displayed in the video, the output frame size is usually smaller than the captured frame, which requires down sampling of the input data. However, at higher zoom levels, up sampling is required to map the ROI to the output video frame. A popular method of electronic zooming is to produce the full color image frame of the total ROA and selecting the ROI window for up/down scaling to fit the output frame. However, this method incur unnecessary computations on color reproduction in the areas out of the ROI and performs scaling on already color interpolated input video frame [4].

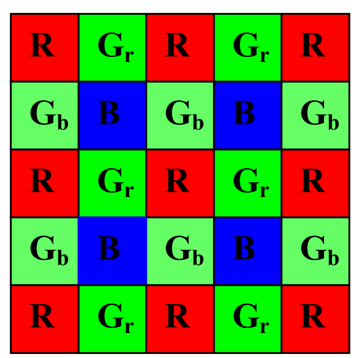


Fig. 1. Color filter array with Bayer pattern.

In single sensor electronic imaging systems, scene color is acquired by sub-sampling in three-color planes to capture color image data simultaneously for red, green and blue color components. Usually this is accomplished by placing a color filter array (CFA) over a 2D sensor array. A type of CFA called the Bayer pattern [5] is shown in Figure 1.

All the Authors of this paper were with Motorola Australian Research Center, Sydney, Australia, when this work was completed. (e-mail: chaminda_w@iprimus.com.au, m.nilsson@telstra.com, serge@tpg.com.au , igor@uow.edu.au)

Contributed Paper
Manuscript received June 9, 2004

This paper presents a method of color interpolation performed on the Bayer pattern, which takes into account the pan tilt and zoom status of the current ROI. This method can be used both for up and down sampling without performing a specific scaling operation. Many methods are described for down scaling in the CFA pattern by sub-sampling [6][7]. However, no methods are described for both up and down sampling of interpolated tri-color data taking into consideration the relative locations of R, G and B components in the captured CFA. This paper describes a method of weighted bilinear interpolation scheme based on relative distance information from the intended interpolated pixel location to the original R, G and B components in the CFA. Performing CFA window based computations for color interpolation is well known from early literature [8] and hence a non-proprietary concept. The output video resolution is simply determined by the choice of the intended interpolated pixel locations rather than performing a specific scaling operation.

II. Camera System Architecture

The block diagram shown in Figure 2 describes the camera system architecture that incorporates the ROI identification and the automatic PTZ feature. The output video frame from the Color Processing module is of constant size 640×256 per PAL field (i.e. 640×512 interleaved). The pan tilt and zoom values are directly used to compute the address of the CFA data window being processed and the weights used in the interpolation scheme. Therefore, no specific scaling step is performed before or after the interpolation or color reconstruction of the video frame.

Image enhancement and noise reduction features are implemented on the interpolated YUV color space within the Image Enhancement module. However, these functions are applied only to the luminance channel (Y). It is also known that sharpness and noise in luminance channel is more visible to human eye. In most image compression schemes, the chrominance channels (U and V) are sub-sampled or smoothed while luminance (Y) is preserved, due to the same reason.

III. PTZ Based Color Interpolation

Color interpolation is based on a weighted average bilinear algorithm. The weights are based on the zoom mode and the distance from the intended interpolated pixel location to the original R, G and B components in the CFA. The start addresses for accessing the CFA data are determined by the current pan and tilt values; whereas the address increment rate is determined by the current zoom value. Zoom values 1,2 and 4 are implemented in the current system. Table 1 shows the relationship between the zoom value and the address increment.

TABLE I
ADDRESS INCREMENT AS A FUNCTION OF ZOOM VALUE

Zoom value	Address increment
1	Twice after each pixel
2	Once after each pixel
4	Once after two pixels

The method of address increment shown in Table 1 together with adaptive weight assignment avoids the usual scale factors associated with conventional scaling. Such scale factors are typically floating point numbers and also involve division operations, which are non-amicable for hardware implementation. The pan (P) and tilt (T) values are also constrained by the current zoom (Z) value as shown in the following equations (1) and (2).

$$0 \le P < 1280 \times \left(1 - \frac{1}{Z}\right) \quad (1)$$

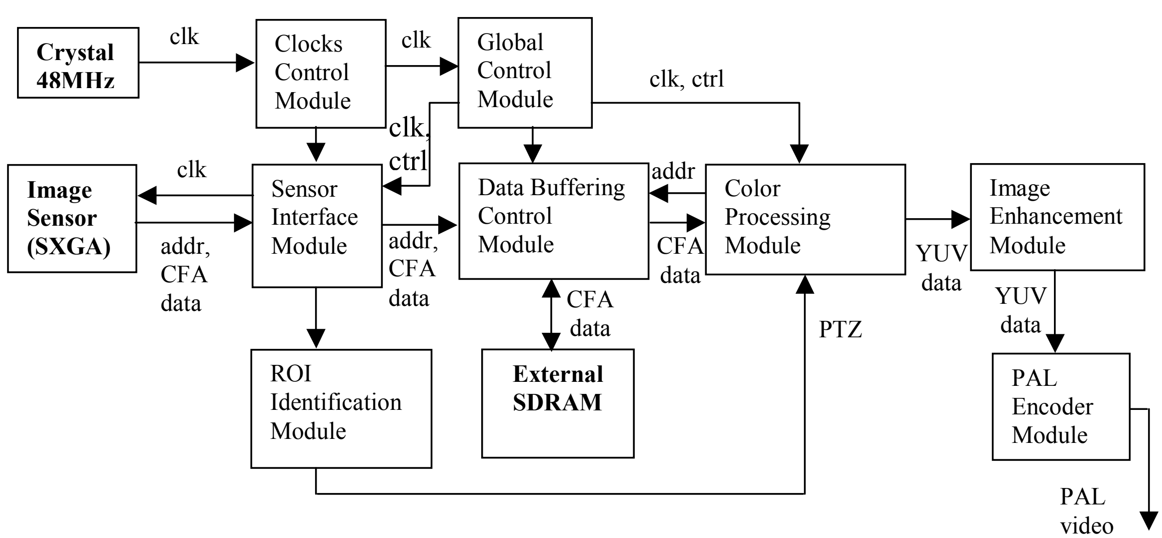


Fig. 2. Camera system architecture block diagram.

$$0 \le T < 1024 \times \left(1 - \frac{1}{Z}\right) \quad (2)$$

The processing CFA data window is always set to be 4×4 regardless of the zoom value. This is also a hardware friendly feature, since only the most demanding process determines the hardware resources. The following sub-sections describe the allocation of weights for performing weighted bilinear interpolation on the CFA data for reconstructing R, G and B color data for each output pixel. The CFA pattern considered is the Bayer pattern shown in Figure 1.

A. Zoomx1 mode

In the zoomx1 mode, CFA frame of size 1280 x 1024 is interpolated to produce a full color frame of size 640 x 512. Therefore, when producing the required 640 (x 3 for R/G/B) values per video line, the column address is advanced by 2 each time, traversing the entire captured data (i.e. 1280 values). Due to this jump of 2, the CFA pattern considered for interpolation remains the same since Bayer pattern is periodic after each two rows and columns.

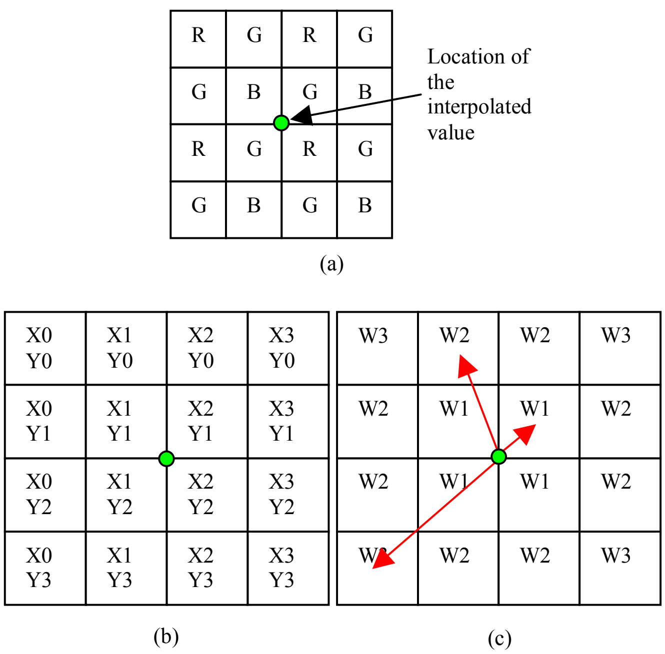


Fig. 3. CFA window for processing in Zoomx1 mode: (a) CFA pattern, (b) data positions and (c) distance based weight distribution.

Assuming square pixels, it is possible to compute the distance from the interpolated pixel location to each weight category pixel (see Figure 3). Let the distances be d_1, d_2, d_3 for weight categories W1, W2 and W3. If a pixel width is taken as 1 unit, $d_1 = 0.707$, $d_2 = 1.581$ and $d_3 = 2.121$. In zoomx1 mode, the weights are assigned as shown in equation (3) to give higher priority to nearby pixels.

$$W \propto \frac{1}{d^4} \quad (3)$$

In this case, W1=0.95, W2=0.035 and W3=0.015 is obtained after normalization to 1. Therefore, W2 and W3 pixel positions are disregarded in zoomx1 mode. The interpolated values are given by equations (4 - 6) for easy hardware implementation.

R = X2Y2; (4)
G = (X1Y2 + X2Y1)/2; (5)
B = X1Y1; (6)

B. Zoomx2 mode

In the zoom2 mode, CFA frame of size 640 x 512 is interpolated to produce a full color frame of size 640 x 512. Therefore, when producing the required 640 (x 3 for R/G/B) values per video line, the column address is advanced by 1 each time, traversing half of the captured data (i.e. 640 values) per line. Therefore, the CFA pattern considered for interpolation, changes (i.e. toggles) according to the row and column number as shown in Table 2.

TABLE II
POSSIBLE CFA PATTERN CASES FOR INTERPOLATION IN ZOOMX2 MODE

Case	REM[row/2]*	REM[col/2]*	Bayer pattern
1	0	0	RG/GB
2	0	1	GR/BG
3	1	0	GB/RG
4	1	1	BG/GR

* Note that REM[.] means the integer remainder resulting from the computation.

Case 1: Identical to Zoomx1 mode, however, in zoomx2 mode, the weights are assigned as shown in equation (7).

$$W \propto \frac{1}{d^2} \quad (7)$$

In this case, W1=0.76, W2=0.15 and W3=0.09 is obtained after normalization. Since the values of W2 and W3 are still significantly smaller than W1, if any color channel contains a single W1 location, this is taken as the interpolated value. This results in equations (4 - 6) for Case 1.

Case 2: For this case, the interpolated values are given by,

R = X1Y2; (8)
G = (X1Y1 + X2Y2)/2; (9)
B = X2Y1; (10)

Case 3: For this case, the interpolated values are given by,

R = X2Y1; (11)
G = (X1Y1 + X2Y2)/2; (12)
B = X1Y2; (13)

Case 4: For this case, the interpolated values are given by,

R = X1Y1; (14)
G = (X1Y2 + X2Y1)/2; (15)
B = X2Y2; (16)

C. Zoomx4 mode

In the Zoomx4 mode, CFA frame of size 320 x 256 is interpolated to produce a full color frame of size 640 x 512. Therefore, when producing the required 640 (x 3 for R/G/B) values per video line, the column address is advanced by 1 every second column count, traversing quarter of the captured data (i.e. 320 values) per line. Therefore, the CFA pattern considered for interpolation changes (i.e. toggles) according to the row and the column number as shown in Table 3.

TABLE III
POSSIBLE CFA PATTERN CASES FOR INTERPOLATION IN ZOOMX4 MODE

Case	REM[row/2]	REM[col/2]	Bayer pattern
1	0	0	RG/GB(1)
2	0	1	RG/GB(2)
3	0	2	GR/BG(1)
4	0	3	GR/BG(2)
5	1	0	RG/GB(3)
6	1	1	RG/GB(4)
7	1	2	GR/BG(3)
8	1	3	GR/BG(4)
9	2	0	GB/RG(1)
10	2	1	GB/RG(2)
11	2	2	BG/GR(1)
12	2	3	BG/GR(2)
13	3	0	GB/RG(3)
14	3	1	GB/RG(4)
15	3	2	BG/GR(3)
16	3	3	BG/GR(4)

Notice that the same CFA data window is now used to interpolate 4 pixel values. If relative distance measures are not used, all the 4 pixels will produce the same result, which will lead to blocking or jagged edges as seen on many digital zooming schemes. For brevity, only the RG/GB CFA window is considered to illustrate the effect of weight assignment. This can be extrapolated to the other 3 types of CFA patterns using

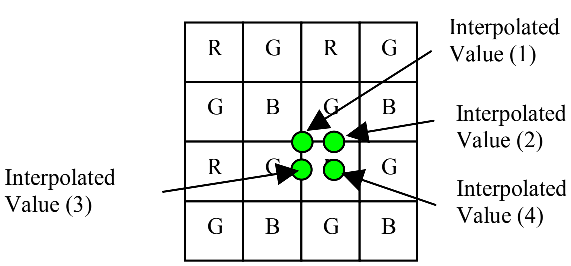


Fig. 4. One of the CFA windows for processing in Zoomx4 mode.

the description on zoomx2 mode. The interpolated value position indicated within brackets near the CFA pattern is with reference to Fig. 4.

With regards to RG/GB pattern, Cases 1, 2, 5 and 6 are considered.

Case 1: Identical to Zoomx1 mode, however, in zoomx4 mode, the weights are assigned as shown in equation (17).

$$W \propto \frac{1}{d} \tag{17}$$

In this case, W1=0.56, W2=0.25 and W3=0.19 is obtained after normalization. Now, the weights W2 and W3 are significant with regards to W1. Therefore, the interpolated values are given by,

G = (3*X2Y1 + 3*X1Y2 + X1Y0 + X2Y3)/8; (18)
(i.e. W1=0.75, W2=0.25, W3=0.00)

R = (5.5*X2Y2 + X2Y0 + 0.5*X0Y0 + X0Y2)/8; (19)
(i.e. W1=0.69, W2=0.25, W3=0.06)

B = (5.5*X1Y1 + X1Y3 + X3Y1 + 0.5*X3Y3)/8; (20)
(i.e. W1=0.69, W2=0.25, W3=0.06)

It can be seen that the hardware amicable equations slightly depart from the computed weight values; however, this is required to optimize the required hardware resources.

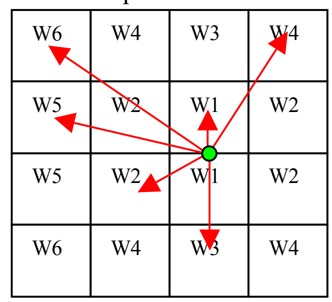


Fig. 5. Weight distribution for zoomx4 mode case 2.

Case 2: Due to the shift in the position of the interpolated pixel, the weight distribution is altered for this case. The new weight distribution is shown in Figure 5.

Geometric analysis produces the following distance measures for each weight category.

$d_1 = 0.5$, $d_2 = 1.12$, $d_3 = 1.5$, $d_4 = 1.8$, $d_5 = 2.06$ and $d_6 = 2.5$

Using Equation (9) and subsequent normalization produces the weights, W1=0.4, W2=0.18, W3=0.14, W4=0.11, W5=0.09 and W6=0.08. Therefore, the interpolated values are given by,

G = (4*X2Y1 + 1.5*X1Y2 + 1.5*X3Y2+X2Y3)/8; (21)
(i.e. W1=0.5, W2=0.375, W3=0.125)

R = (6*X2Y2 + 2*X2Y0)/8; (22)
(i.e. W1=0.75, W3=0.25) from W1, W3, W5 and W6

B = (3*X1Y1 + X1Y3 + 3*X3Y1 + X3Y3)/8; (23)
(i.e. W2=0.75, W4=0.25) from only W2 and W4

It should be noted that each color channel represents not all the weight categories. Therefore, it is necessary to re-normalize the weight categories that are within each color channel disregarding the weight categories not represented, as in the case of R and B.

Case 5: Due to the shift in the position of the interpolated pixel, the weight distribution is altered for this case. The new weight distribution is shown in Figure 6.

Although the weight distribution has been changed, the distances to weight categories and hence the weights are as same as in case 2. Therefore, the interpolated values are given by,

G = (1.5*X2Y1 + 4*X1Y2 + 1.5*X2Y3 + X3Y2)/8; (24)
(i.e. W1=0.5, W2=0.375, W3=0.125)

R = (6*X2Y2 + 2*X0Y2)/8; (25)

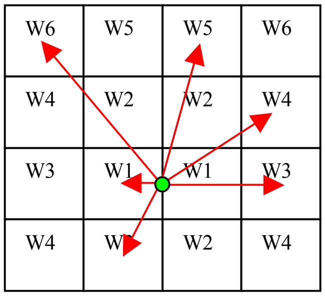


Fig. 6. Weight distribution for zoomx4 mode case 5.

(i.e. W1=0.75, W3=0.25) from W1, W3, W5 and W6

B = (3*X1Y1 + 3*X1Y3 + X3Y1 + X3Y3)/8; (26)

(i.e. W2=0.75, W4=0.25) from only W2 and W4

Case 6: Due to the shift in the position of the interpolated

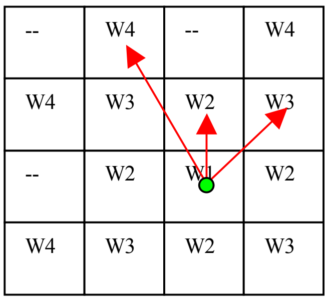


Fig. 7. Weight distribution for zoomx4 mode case 6.

pixel, the weight distribution is altered for this case. The new weight distribution is shown in Figure 7.

This is a special case since the interpolated pixel is located on one of the captured CFA data. Therefore, color channel value is generated from this pixel and all other same color values are disregarded in weight distribution setting. Geometric analysis produces the following distance measures for each weight category.

$d_1 = 0.0$, $d_2 = 1.0$, $d_3 = 1.4$ and $d_4 = 2.2$

Since $d_1 = 0.0$ it is not used in the weight computation. Using Equation (9) and subsequent normalization produces the weights, W2=0.46, W3=0.33, W4=0.21. Therefore, the interpolated values are given by,

G = (2*X2Y1 + 2*X1Y2 + 2*X3Y2 + 2*X2Y3)/8; (27)

(i.e. W2=1.0) from only W2 and W4

R = (8*X2Y2)/8; (28)

(i.e. W1=1.0) from only W1

B = (2*X1Y1 + 2*X1Y3 + 2*X3Y1 + 2*X3Y3)/8; (29)

(i.e. W3=1.0) from only W3

If the interpolated pixel is located on the CFA pixel or a particular color channel is significantly represented only by a particular weight category, it is necessary to allocate a weight of 1.0 for that particular category, as seen in this case.

This zooming strategy can be used for zoomx8 and higher zooming levels. In zoomx8 mode, the weights are assigned as shown in equation (30).

$$W \propto \frac{1}{\sqrt{d}} \quad (30)$$

This method can also be used in conjunction with gradient detection, to avoid using pixel values across edges. However, additional computations are required in this case, hence increasing the hardware resources needed.

IV. Color Correction

Color correction is performed using an algorithm developed by the authors, which is described in detail in [9]. The Green channel is unaltered. Red and Blue channels are corrected using the following formulae.

$$R_c = R + 0.5\times(\overline{R} - \overline{G}) \quad (31)$$

$$B_c = B + 0.5\times(\overline{B} - \overline{G}) \quad (32)$$

$$G_c = G \quad (33)$$

The $\overline{R}, \overline{G}, \overline{B}$ values represent the average channel value computed for each CFA window of size 4×4. This computation is performed during the color interpolation, disregarding the weight values. The resulting values are clipped to be within the range [0…255].

V. Color Space Conversion

Color space conversion from RGB to YUV is performed using the following formulae. The coefficients are adjusted to be hardware friendly, without compromising visual quality.

$$Y = 0.5\times R_c + 1.25\times G_c + 0.125\times B_c \quad \text{range [0…511]} \quad (34)$$

$$U = B_c - 0.5\times Y \quad \text{range [-256 …255]} \quad (35)$$

$$V = R_c - 0.5\times Y \quad \text{range [-256 …255]} \quad (36)$$

VI. Sharpness Enhancement

Methods of sharpness enhancement found in literature [10][11] are in the context of compression algorithms (e.g. JPEG), whereas a few deal directly on the luminance (Y) and chrominance (U/V) signals [12]. The constrains on the current implementation are outlined below:

(1) Sharpening and noise reduction should be implemented on the YUV color space

(2) Algorithms should be line based to avoid using temporary storage space

(3) Algorithm filter taps should be minimized to preserve as many output pixels, since each extra tap eliminates one valid pixel.

Under these constraints, a simple non-linear sharpening filter was implemented. The filter has the following properties:

Filter taps: S_{-2}, S_{-1}, S_0, S_{+1}, S_{+2}

$$Y' = |S_1 - S_{-1}|$$

Filter coefficients: If ($Y' > 20$) then {-0.25, -0.25, 2.00, -0.25, -0.25} else {0.00, 0.00, 1.00, 0.00, 0.00}.

Although the sharpening is performed only on the Y channel, it is important to store the corresponding U and V values to avoid color artifacts. This method performs well for natural images to enhance the appearance of sharpness to the

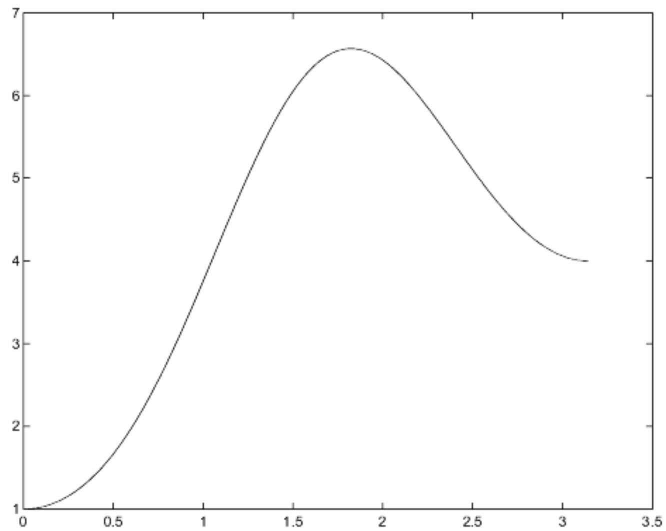


Fig. 8. Frequency response of the implemented filter.

human viewer without significantly increasing the noise level. However, the non-linear processing may inflict some artifacts especially near high frequency thin lines (above 5MHz). Figure 8 shows the frequency response of the implemented filter.

By enhancing the gain of the spatial frequencies in the vicinity of the peak response of the human Visual System (HVS), it is possible to make the output image "appear" sharper and with higher contrast. However, the distance of the viewer from the image display device also dictates this effect.

VII. NOISE REDUCTION

The implemented noise a suppression algorithm is based on anisotropic diffusion [13]. Anisotropic diffusion is performed along each video line, for Y channel only. The main reason for this is to minimize the data buffers needed for storing intermediate data between iterations. Only 3 iterations are performed using intermediate storage. Although only Y channel is processed, corresponding U and V values should also be saved, to avoid color shifts at the output image.

A. Anisotropic Diffusion Algorithm

Diffusion algorithms remove noise from an image by modifying the image via a partial differential equation (PDE). Anisotropic diffusion stops the diffusion across edges preserving the edge sharpness in the image [13]. Much research has been conducted on the characteristics and behavior of anisotropic diffusion on images [14]. The following describes the implementation of a line based anisotropic diffusion algorithm for noise reduction while preserving the edge sharpness. The correction weights for each pixel Y value are based on its immediate horizontal gradient. Consider the following scenario:

Y_{t-2}, Y_{t-1}, Y_t

The $Y_{t'}$ which replaces the Y_{t-1} is computed as follows:

$$\nabla_1 = Y_{t-2} - Y_{t-1} \tag{37}$$

TABLE IV
CORRECTION WEIGHT ASSIGNMENT BASED ON ABSOLUTE GRADIENT VALUE

$\lvert\nabla_1\rvert$	$C_{\nabla 1}$	$\lvert\nabla_2\rvert$	$C_{\nabla 2}$
>32	256	>32	256
16 - 32	4	16 - 32	4
8 - 16	2	8 - 16	2
0 - 8	1	0 - 8	1

$$\nabla_2 = Y_t - Y_{t-1} \tag{38}$$

The correction weights are produced as shown in Table 4.

$$Y_{t'} = Y_{t-1} + \left(\frac{\nabla_1}{C_{\nabla 1}} + \frac{\nabla_2}{C_{\nabla 2}} \right) \tag{39}$$

The resulting $Y_{t'}$ value is clipped to be within the range [0…511]. This algorithm is particularly effective in high noise conditions (e.g. at low light).

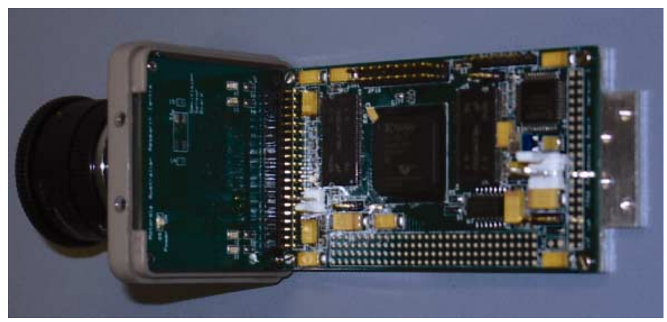


Fig. 9. Camera platform incorporating automatic PTZ and image enhancement.

VIII. RESULTS AND DISCUSSION

The proposed algorithms for the camera system were implemented on a Xilinx Virtex II Field Programmable Gate Array (FPGA) based platform as shown in Figure 9.

The following camera testing procedures were followed to test the camera platform including the lens system and the IR filter. The test procedures are documented in [15]. The test results are subjected to change if a different lens or a different IR filter was used. The following aspects were measured for benchmarking:

(1) Resolution of the image at normal, zoom2 and zoom4 modes.
(2) Sharpness of the picture at normal, zoom2 and zoom4 modes.
(3) Noise performance at normal, zoom2 and zoom4 modes.

The developed camera platform was compared with Samsung SDC-450 camera with a comparable lens. The testing equipment (shown in Figure 10) consisted of,

(1) Tektronics WFM91 waveform/vector scope field unit with monitor

(2) Tektronics TG2 test pattern generator
(3) GrabBee USB based video capture device connected to PC
(4) CCTV Labs Test Chart V.1.3 prepared by Vlado Damjanovski

Fig. 10. Test Equipment (pattern generator and a vectorscope).

(5) SONY Video Monitor PVM-20M4A
(6) TENMA 72-6693 Light Meter

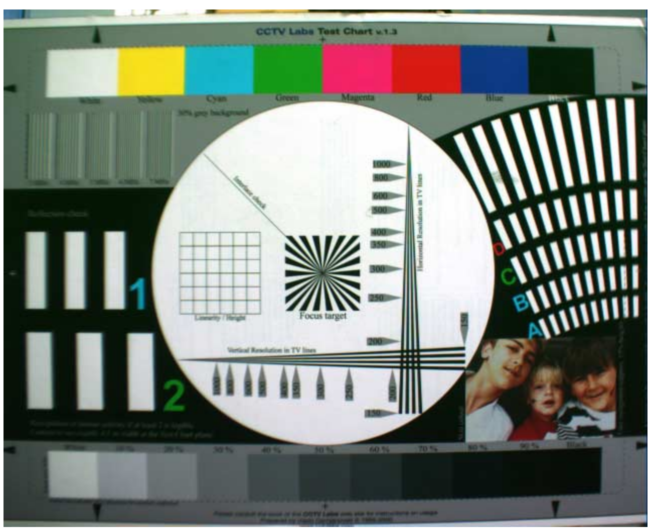


Fig. 11. CMOS Camera Platform – Zoomx1 mode

TABLE V
RESOLUTION COMPARISON

Mode	CMOS Camera Platform	Samsung SDC-450
Horiz. (Zoomx1)	300 TV lines	350 TV lines
Vert. (Zoomx1)	300 TV lines	350 TV lines
Horiz. (Zoomx2)	600 TV lines	N/A
Vert. (Zoomx2)	600 TV lines	N/A
Horiz. (Zoomx4)	800 TV lines	N/A
Vert. (Zoomx4)	800 TV lines	N/A

A. Equipment Set-up

The video monitor was adjusted to the standard conditions (e.g. brightness, color, phase etc.) using the Tektronics TG2 test pattern generator. The room lighting (fluorescent) level was measured using TENMA 72-6693 Light Meter, and was recorded to be 187Lux. The color temperature of the room lighting was slightly unbalanced producing slightly more reddish than bluish. Since it is cumbersome to balance the lighting levels, the fact that the lighting was biased towards red should be taken in to account when interpreting the test results. The cameras were set-up directly in front of the test chart at 0.8 meters away. The optical zoom was adjusted so that the chart fully occupies the total field of view. The cameras were untouched after set-up.

B. Resolution:

Table 5 shows objective camera testing results in terms of resolution and presents an objective comparison between the CMOS camera platform and CCD based Samsung camera.

The images in Figures 11 – 15 depict the resolution comparison subjected to the limitations of the frame grabber.

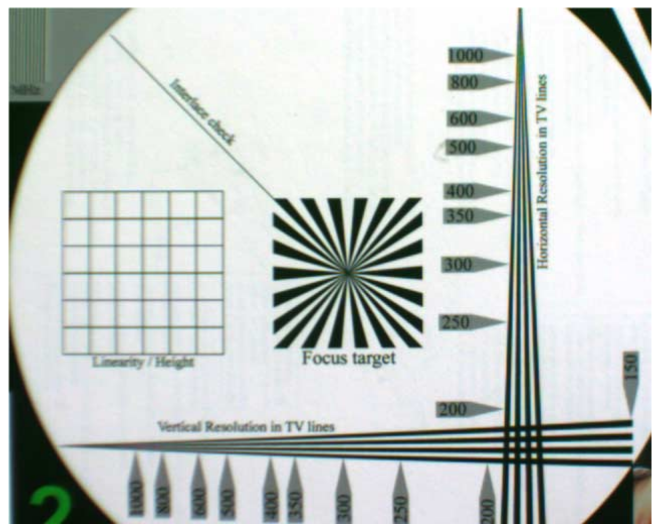


Fig. 12. CMOS Camera Platform – Zoomx2 mode

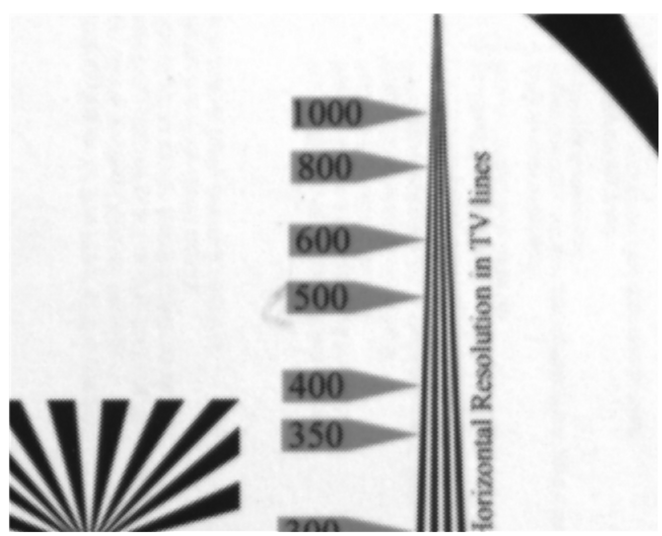


Fig. 13. CMOS Camera Platform – Zoomx4 mode (horiz. Res.)

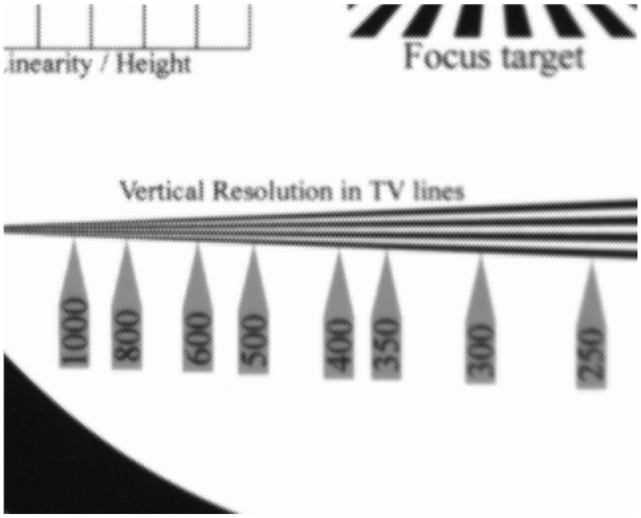


Fig. 14. CMOS Camera Platform – Zoomx4 mode (vert. Res.)

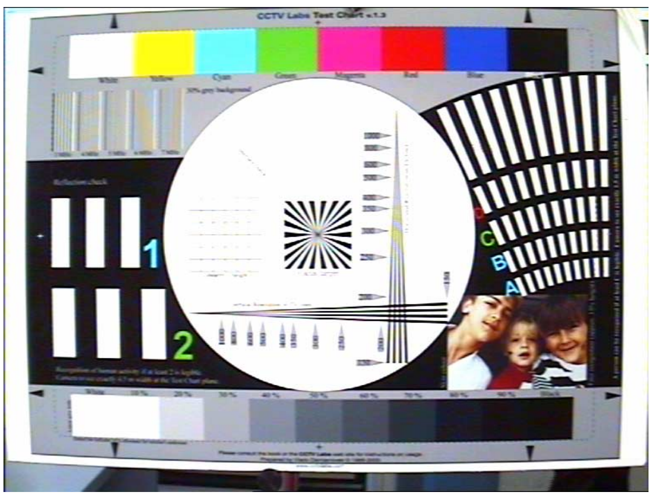


Fig. 15. CCD camera – full frame mode is only available

C. Sharpness

Table 6 shows objective camera testing results in terms of sharpness and presents an objective comparison between the

TABLE VI
SHARPNESS COMPARISON

Mode	CMOS Camera Platform	Samsung SDC-450
Zoomx1	4 MHz	3 MHz
Zoomx2	6 MHz	N/A
Zoomx4	7 MHz	N/A

CMOS camera platform and CCD based Samsung camera.

The figures 16 - 18 depict the sharpness comparison subjected to the limitations of the frame grabber.

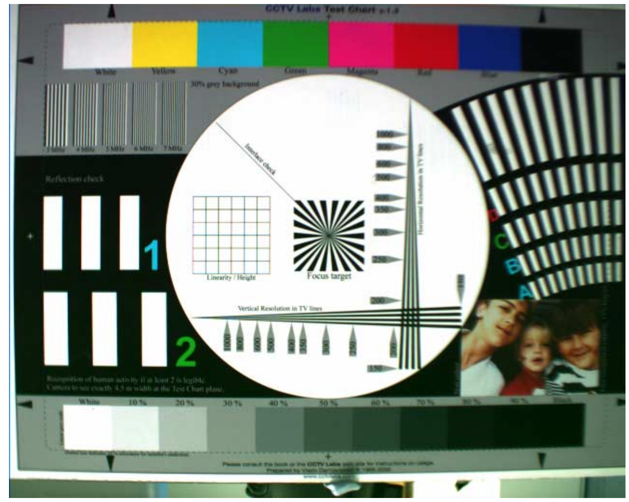


Fig. 16. CMOS Camera Platform – Zoomx1 mode (up to 4MHz lines visible on TV)

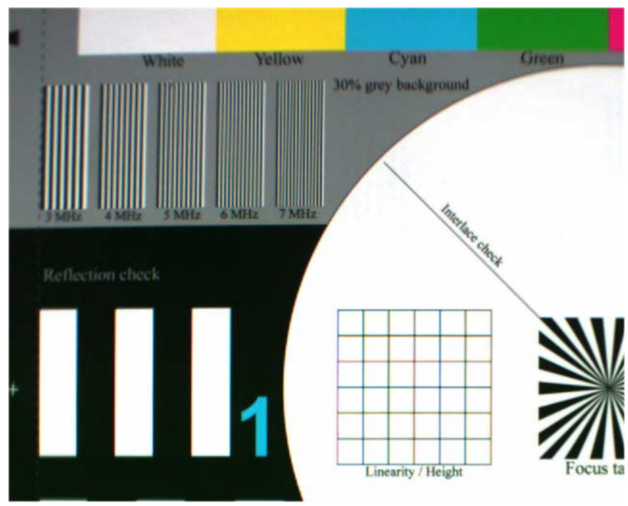


Fig. 17. CMOS Camera Platform – Zoomx2 mode (up to 6MHz lines visible on TV)

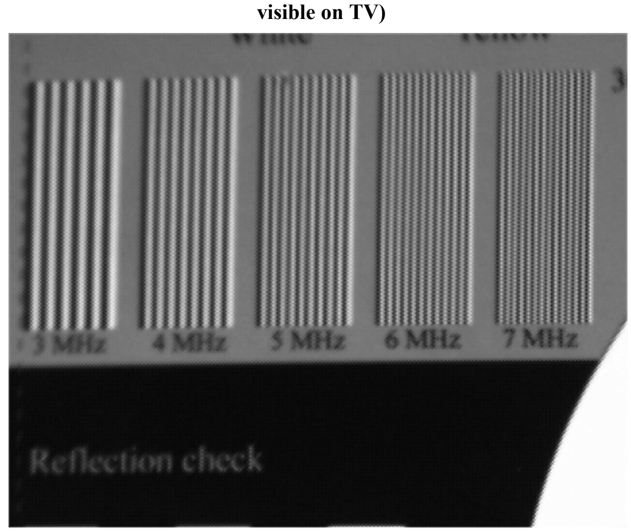


Fig. 18. CMOS Camera Platform – Zoomx4 mode (up to 7MHz lines visible on TV)

It should be noted that when digital zoom modes were used, the resolution of the display TV is not a limiting factor, since at zoomed view 7MHz do not represent 7MHz sharpness on the display. However, since no optical zooming was used, in terms of what the camera can see and interpolate at zoomx4 represent the sharpness quality equivalent to 7MHz.

D. Noise performance:

The noise was measured using the gray scale levels on a selected PAL video line. Figure 19 –20 shows the captured waveforms for CMOS camera platform and CCD camera.

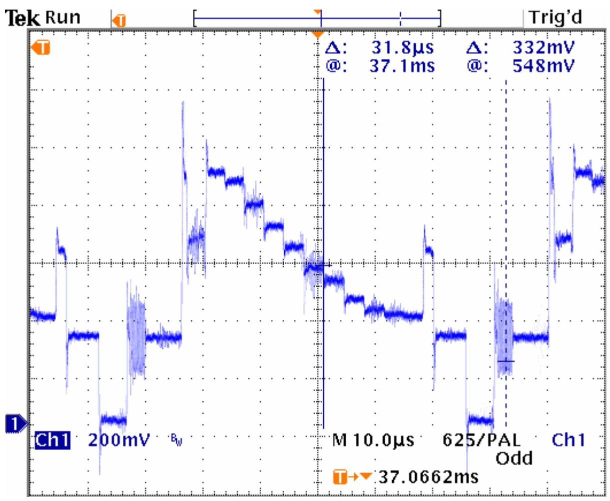


Fig. 19. CMOS Camera Platform – Gray levels at fluorescent lighting (200 mV/div scale)

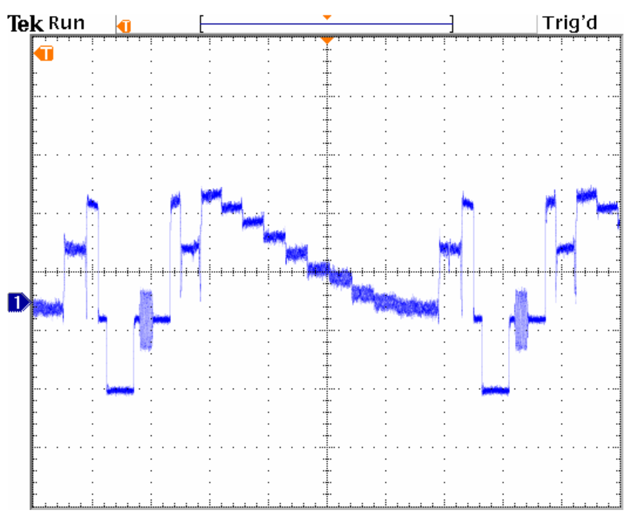


Fig. 20. CCD camera gray levels in fluorescent lighting (500 mV/div scale)

TABLE VII
NOISE LEVEL COMPARISON

	CMOS Camera Platform	Samsung SDC-450
Signal to Noise Ratio (SNR)	29 dB	24 dB

Table 7 shows objective camera testing results in terms of noise performance and presents an objective comparison between the CMOS camera platform and CCD based Samsung camera at 187Lux luminance level.

The main reason for higher noise level in CCD camera is due to the attempt at over saturating colors. At gray levels this introduces more noise and at low light levels high color noise is observed. CMOS camera platform uses the color correction algorithm described in Section 4 and also uses noise reduction method described in Section 7. They collectively contribute towards lower output noise although the input noise level is higher for CMOS compared to CCD sensors.

IX. CONCLUSION

Development of a digital zoom camera with image sharpening and noise reduction is described in this paper. Digital zooming is accomplished using a weighted bilinear interpolation scheme based on relative distance information from the intended interpolated pixel location to the original R, G and B components in the CFA. The output video resolution is simply determined by the choice of the intended interpolated pixel locations rather than performing a specific scaling operation. Image sharpening is performed using a non-linear sharpening filter based on the luminance (Y) gradient and a filter with mid-frequency amplification. Noise reduction is accomplished using a line based anisotropic diffusion algorithm. The proposed system is implemented using an FPGA based CMOS camera platform. Camera testing indicates superior performance in resolution, sharpness and noise suppression against standard CCTV cameras in zoomx2 and zoomx4 modes.

ACKNOWLEDGMENT

The authors would like to thank Mr. Ollencio DeSouza of Chubb Security Australia for his input to this paper and for providing camera-testing equipment for gathering results.

REFERENCES

[1] G. Thiel, “Automatic CCTV surveillance – Towards the Virtual Guard”, IEEE AES Systems Magazine, July issue, pp 3 – 9, 2000.

[2] M. Nicolescu and G. Medioni, “Electronic Pan-Tilt-Zoom: A solution for intelligent room systems”, Proc. of IEEE International Conference on Multimedia and Expo (ICME 2000), Vol. 3, 30 July-2 Aug., pp 1581 –1584, 2000.

[3] G. Meyanants, B. Dierickx, A. Alaerts, D. Uwaerts, S. Cos, D. Scheffer and S. Noble, “A 35 mm 13.89 Million pixel CMOS active pixel image sensor”, Proc. of IEEE

workshop on CCD & AIS, May 15 – 17, Elmau, Germany, 2003.

[4] U.S. Patent Application No.2003/0007082 A1, S. Watanabe, "Digital camera with electronic zooming function", Casio Computer Co. Ltd.

[5] U.S. Patent No. 3,971,065, Bayer, Eastman Kodak Company

[6] U.S. Patent Application No. 2003/0122937 A1, M. Guarnera et al, "Method for processing digital CFA images, particularly for motion and still imaging"

[7] U.S. Patent Application No. 2003/0020819 A1, H. Fukuda, "Image pickup apparatus".

[8] K.A. Parulski, "Color filters and processing alternatives for one-chip cameras", IEEE Trans. on Electron Devices, Vol. ED-32(8), pp 1381, 1985

[9] I. Kharitonenko, S. Twelves and C. Weerasinghe, "Suppression of noise amplification during color correction", IEEE Trans. Consumer Electronics, Vol. 48 (2), pp. 229 – 233, 2002

[10] M.Fischer, J.L. Paredes, G.R. Arce, "Weighted median image sharpeners for the World Wide Web", IEEE Trans. on Image Processing, Vol. 11 (7), pp 717 – 727, 2002

[11] K. Konstantinides, V. Bhaskaran, G. Beretta, "Image sharpening in the JPEG domain", IEEE Trans. on Image Processing, Vol. 8 (6), pp 874 – 878, 1999

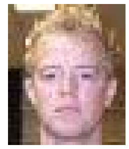
[12] S. J. Huang, "Adaptive noise reduction and image sharpening for digital video compression", Proc. of IEEE International Conference on Systems, Man, and Cybernetics, Vol. 4, pp 3142 – 3147, 12-15 Oct., 1997

[13] Perona P. and Malik J., "Scale-space and edge detection using anisotropic diffusion", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, No. 7, PP 629-639, July 1990.

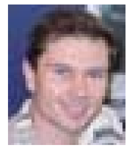
[14] Black M.J. and Marimont D.H., "Robust Anisotropic Diffusion", IEEE Trans. On Image Processing, Vol. 7, No. 3, pp 421-432, March 1998.

[15] DeSouza O., "Objective Testing Techniques", Technical Brief, Chubb Security Australia, 2002.

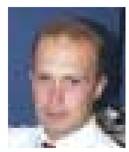
Chaminda Weerasinghe received BE Honors Class 1 with university medal from University of Wollongong, Australia in 1994 and his Ph.D. in image processing from University of Sydney, Australia in 1999. He is a recipient of many academic awards and medals from IEE, IEAust and IESA. Dr. Weerasinghe was with Motorola Australian Research Center (2000 – 2003), as a senior research engineer. He is currently with Toshiba (Australia) Pty. Ltd. (R&D Division). His main research interests are in color image processing; CMOS image sensors, surveillance camera systems, stereoscopic/panoramic video generation/display and raster image processing for printers..



Magnus Nilsson was born in Gothenburg, Sweden in 1976. He received his BSc (electrical engineering) in 2000 from Chalmers University of Technology, Sweden, and his Master of Computer and Information Engineering in 2002 from Griffith University, Australia. He was working as a senior research engineer at Motorola Australian Research Center in Sydney, Australia (2001 – 2003). His research interests include Smart CMOS image sensors and embedded speech and image processing.



Serge Lichman received the M.Sc. degree in computer systems from the Odessa National Polytechnic University in 1988. After graduation he worked in various positions as electronic/software design engineer. He was a Senior Research Engineer at Motorola Australian Research Centre (2001 – 2003). He is currently with NICTA, Sydney, Australia. His research interests are in computer vision including robust motion detection and tracking.



Igor Kharitonenko was born in Odessa, Ukraine. He received the B.S. with honors in electronics engineering and Ph.D. degree from Odessa Polytechnic University in 1985 and 1993 respectively. Dr. Kharitonenko was a Principal Research Engineer at Motorola Australian Research Centre (1995 – 2003) working on technology development for digital cameras and mobile video communicators. Since 1997 till 2000 he was involved in ISO activity on JPEG2000 image compression standard development. He is currently with University of Wollongong. His research interests include machine vision, multimedia content management, intelligent surveillance systems, CMOS image sensor architectures, image and video compression.