AN EDGE DETECTION APPROACH TO DIGITAL IMAGE STABILIZATION BASED ON TRI-STATE ADAPTIVE LINEAR NEURONS

Joon Ki Paik, Yong Chul Park, and Sung Wook Park
Video System Tech. Team, Buchun R&D Center, Semiconductor Business, Samsung Electronics Co.

ABSTRACT

A new approach to digital image stabilization (DIS) for video cameras is proposed. The proposed DIS system is composed of an edge detection unit, a motion detection unit, and a digital zooming unit. The edge detection unit is based on tri-state adaptive linear neurons (ADALINES) [1], the motion detection unit is based on the corresponding binary logical correlation computations instead of multiple-bit pixel-wise subtractions, and the digital zooming unit is based on the approximated bilinear interpolation (ABI) technique.

A motion decision procedure is also proposed for removing unwanted effects on the integrated motion vector. The proposed DIS system is designed manily for the hardware minimization without using additional Micoms in a video camera system. Experimental results show that the proposed system accurately detects and compensates camera motion.

I. INTRODUCTION

In recent days, video cameras for consumer use exhibit certain changes such as: (1) compactness in size, (2) high zooming ratio, and (3) digitalization in video signal processing. (1) and (2) make it necessary for video cameras to have an image stabilizing system [2]-[4], and (3) makes it easy to implement the stabilizing system. A bolck diagram of a digital video camera with a DIS system and a corresponding field memory (FM) is shown in Fig. 1. A DIS system usually has the following three functional units, such as: (1) a pre-processing unit, (2) a motion vector detection unit, and (3) a motion compensation and digital zooming unit. The proposed DIS system uses an edge detection method based on the tri-state ADALINES in the pre-processing unit, a logical correlation computation method in the motion vector detection unit, and the ABI method in the digital zooming unit, as shown in Fig. 2.

A major advantage of the proposed DIS system is that it greatly reduces the computational load in detecting the global motion vector between a pair of images, say the current and the previous image frames, by the use of binary edge images. The probems in using the binary edge images will be discussed and the promising solution will be given in Secs. II and III.

Interpolation, which is used in the digital zooming unit of the proposed DIS system, is an important issue in digital image and video signal processing areas. An efficient way in implementing the digital zooming is also given in Sec. IV. In Sec. V simulation results of the proposed DIS system are given by using two real-time sampled image sequences.

II. EDGE DETECTION UNIT

Before detecting motion vectors in the DIS system, a pre-processing is recommended because it can improve the accuracy in detected motion vectors by removing unwanted factors in the image such as noise, and at the same time it can reduce the computational load by mapping an image usually represented by 8 bits per pixel, to an image represented by less than 8 bits per pixel. Several kinds of pre-processing method were proposed. For example, Band Extract Representative Point (BERP) method, a kind of bandpass filtering method, was used in [3]. While the BERP method can efficiently filter out both the extremely high spatial frequency components such as noise and low frequency componenets such as the flat area in intensity, it requires great computational efforts for obtaining the BERP image, and it still needs more than two bits per pixel to represent the BERP image.

On the other hand, the edge detection method based on the tri-state ADALINES described in this paper have the following advantages:

(1) Advantages in using the tri-state ADALINES edge detection method;

- It uses 3x3 local windows for mapping each edge value, whereas others should use much larger local windows
- It simultaneously detects edges and suppresses noise. Noise suppression does not need additional computation.
- Since it has an adaptive linear neuron structure, the advantages of neural networks can be benefited in this method such as; the high computational power by parallelism [5] and the rotation invariant property of layered neural networks [6].

(2) Further advantages by the edge detection method

- An edge image requires only one bit to represent a pixel

Manuscript received June 7, 1991

intensity value.

-When computing correlations between two edge images, it requires only simple logical operations, while other methods require multiple-bit pixel-wise subtractions.

-Edge information in an image is known to be very useful and efficient to characterize the motion of objects [7].

The proposed tri-state ADALINES edge detection unit includes a mapping block, a matching block, and a decision block, as shown in Fig. 3. The mapping block adaptively converts intensity values of an image pixel (x_i) and its neighbor pixels (z_j) into appropriate tri-state values (v_j), using a mean value. By the use of an appropriate tri-state mapping, noise effect can be reduced without using any additional lowpass filters. The matching block accepts the corresponding tri-state input values to compare the input pattern with several prespecified simple patterns (w_1-w_4) using a set of tri-state ADALINES, and outputs matching signals. Theoretical details of the pattern matching ability of the tri-state methods, together with the way of mapping tri-state input values, was given in [8]. Using several different matching signals from each tri-state ADALINE, the decision block (OR) finally decides whether the pixel of interest is an edge or not. In demonstrating the performance of the proposed edge detection method, Figs. 4(a) and (b) show the original lady image and the corresponding edge image, respectively. Figs. 4(c) and (d) also show the edge detection result with additive noise (SNR = 10dB).

III. MOTION DETECTION UNIT

After detecting edges, a motion vector is to be determined between two dynamic image sequences. One major difference in detecting motion vectors for the DIS system and other applications is that the DIS system requires only one global motion vector from the two image sequences, while others, for example, motion compensated image coding, requires many motion vectors corresponding to each pixel. Furthermore, the DIS system concerns about the motion of the entire image frame, not about the motion of the object in the image. For this purpose, the motion detection unit is divided by two parts; *motion estimation part* and *motion decision part.*

A. Motion estimation part.

In the motion estimation part, we estimate four motion vectors from the correspoding four areas in the image. A typical method to choose the four motion estimation areas (MEAs) is shown in Fig. 5. As shown in Fig. 5, each MEA is placed near the boundary of the image, in order to avoid detecting the motion of the objects.

The CCD shown in Fig. 1 produces two different signals, say s1 and s2, interchangeably, for efficient color signal processing. Four MEAs are obtained by converting the corresponding areas in two adjacent even (or odd) fields of s1 (or s2) images into binary edge images. More specifically, the horizontal resolution of each MEA is reduced by half due to selecting one of each s1 and s2, and the vertical resolution of each MEA is also reduced by half due to selecting one of even and odd fields of the image. The advantage of using the half resolution MEAs is that it can greatly reduce the computational load, which affects the hardware complexity. On the other hand, use of the half resolution MEAs degrades the estimated motion vector quality, i.e., the maximum error in the estimated motion vector is 1 pixel, while the maximum error in the vector is 0.5 pixel when the full resolution MEA is used. Although the error in the motion vector directly degrades the performance of the DIS system, it can be adjusted in the motion decision part described in the following subsection. The quality of the motion vector can be further improved by using the luminance signal, which is the sum of s1 and s2 signals, in the horizontal direction and by using line linterpolation in the vertical direction.

The first even field of s1 image is called the reference field, and the next even field of s1 image the compared field. Block matching type motion detection method is used to estimate four motion vectors from the corresponding MEAs. The proposed motion estimation part is shown in Fig. 6. Each MEA in the previous image field is latched in the reference field block by using shift registers. And then corresponding blocks in the current image field is sequentially latched in the compared field block. Each compared field block is then compared with the reference field block by summing absolute differences of entire pixels in the block. Since each MEA in both image fields have been converted to binary edge images in the pre-processing stage, pixel-wise subtractions are not needed. Instead, we only need to check if the corresponding pairs of pixels in two image fields have the same 1's or 0's. Then a position of the bolck in the compared field which gives the maximum number of matchings becomes an estimated motion vector. In order to estimate the four motion vector in the corresponding four MEAs in real-time, the MEAs data control block and the demultiplexer is used. The motion address generation block in Fig. 6sequentially generates addressed for candidates of motion vector, and then the block sequentially overwrites the position which gives the maximum correlation. After scanning an entire MEA, a motion vector is obtained from the motion vector estimation block.

B. Motion decision part.

Let (x_i, y_i), i=1,...,4, be four motion vectors obtained in the motion estimation part. Four (x_i, y_i)'s are ideally identical, but the experiment reveals that the motion vectors tend to differ by the following reasons: (1) low contrast, (2) noise in the image data, (3) zooming and panning effects, and (4) motion of an object in an MEA.

Problems (1) and (2) usually result in unpredictable effects, and can be compensated in the edge detection unit by the tri-state mapping technique. On the other hand, (3) and (4) can be measured by an appropriate technique. For example, a technique for measuring (3) has been proposed in the literature [9]. In addition to the above mentioned reasons, many unexpected situation can occur, and they significantly degrade the performance of the DIS system.

In order to obtain the desired motion vector from the four (x_i,y_i)'s which suffer from (3) and (4), the following motion decision technique is used.

Motion decision procedure:

(1) Compute the *isolativity* $I(x)_i$ and $I(y)_i$ by

$$I_x(i) = | x_i - m_x | \text{ and } I_y(i) = | y_i - m_y |, \; i=1, ..., 4,$$

where

$$m_x = (\sum_{i=1}^{4} x_i)/4 \text{ and } m_y = (\sum_{i=1}^{4} y_i)/4.$$

And then compute the *isolation weights* $a_x(i)$ and $a_y(i)$, such as

$$a_x(i) \propto 1/I_x(i) \text{ and } a_y(i) \propto 1/I_y(i).$$

(2) Compute the *stability weights* $s_x(i)$ and $s_y(i)$ such as

$$s_x(i) \propto 1/| x_i - x_{old} | \text{ and } s_y(i) \propto 1/| y_i - y_{old} |,$$

where (x_{old}, y_{old}) is the motion vector computed from the previous pair of images.

Fig. 7 shows the block diagram of the motion decision part. Using four (x_i,y_i)'s, isolation and stability weights are computed first, and then the integrated motion vector is obtained by

$$x = (a_x(1)s_x(1)x_1 + ... + a_x(4)s_x(4)x_4)/(a_x(1)s_x(1) + ... + a_x(4)s_x(4))$$

and

$$y = (a_y(1)s_y(1)y_1 + ... + a_y(4)s_y(4)y_4)/(a_y(1)s_y(1) + ... + a_y(4)s_y(4)).$$

An example for the characteristcs of the isolation and stability weights is shown in Fig. 8. The isolation weights decreases as the isolativity increases, and it becomes zero when the isolativity is larger than some predefined value. On the other hand the stability weights decreases but it is not equal to zero, in order to avoid rejecting a good motion vector.

IV. DIGITAL ZOOMING UNIT

In order to compensate the motion of the compared image, we have moved it along the inverse direction of the motion vector obtained in the motion detection unit. With the assumption that we do not have any information outside the boundary of images, we expand both the reference and compared images by *digital image interpolation technique*.

Various techniques have been proposed for digital image interpolation. Nearest neighbor interpolation (NNI) or replication method is one of the most simplest type technique, where the interpolated pixel takes the intensity value of the nearest pixel [10]. While the NNI method can easily be implemented in hardware, it produces the mosaic effects near the slant edge areas. On the other hand, many sophisticated interpolation techniques, such as higher order polynomial interpolation methods [11] and an algorithm which uses rank order filters for preserving edges [12], was proven to give high quality interpolated images, but they are not suitable for real-time hardware implementation due to their computational complexity.

Typical range of the zooming ratio for the DIS system is 1.1 to 1.5. Since the required zooming ratio is not high, the bilinear interpolation method, which is ranked in the middle of NNI and more sophisticated methods in [11] and [12], is suitable for the DIS system. By this reason, we propose the approximated bilinear interpolation (ABI) method, which realizes discrete zooming ratio of multiple of 0.25 with particularly simple hardwares.

The block diagram of the digital zooming unit is shown in Fig. 9. Once a motion vector is obtained in the motion detection unit, motion compensation should be performed by shifting the entire image field in the inverse direction of the motion vector. In the proposed DIS system, motion compensation and ABI are performed in the zooming unit. In other words, using the motion vector obtained by the motion detection unit and the given zooming ratio (1.25 in the proposed system), the addresses to be read from the field memory (FM) is generated. Image data according to the address are read and zoomed by the ABI method. Horizontal zooming in the ABI block in Fig. 9 is performed for s1 and s2 signals separately for preserving color information, as shown in Fig. 10. For the s1 signal, for example, horizontal zooming receives four inputs, say, b_1, b_2, b_3, b_4, then produces the output z such as

$$z = ((b_1+b_2)/2 + (b_3+b_4)/2)/2,$$

where each division by 2 truncates the floating part in order to avoid increasing the hardware. The maximum truncation error is 0.5 in pixel intensity, which is negligible in the 8 bit pixel intensity range (0 - 255).

In order to produce s_{11}' in Fig. 10, $b_1=b_2=b_3=b_4=s_{11}$. Similarly, for s_{12}', $b_1=s_{11}$ and $b_2=b_3=b_4=s_{12}$; for s_{13}', $b_1=b_2=s_{12}$

and $b_3=b_4=s_{13}$; for s_{14}', $b_1=b_2=b_3=s_{13}$ and $b_4=s_{14}$; and for s_{15}', $b_1=b_2=b_3=b_4=s_{14}$. Vertical zooming is performed in the same manner using a 2H delay. The simulation result of the proposed ABI method is compared with that of Fig. 11(a)-(d). Fig. 11(a) is the original cameraman image. Fig. 11(b) is the 4 times zoomed image by the original bilinear interpolation method. The zoomed image by the NNI method, shown in Fig. 11(c), exhibits the mosaic effects. On the other hand, the zoomed image by the proposed ABI method, shown in Fig. 11(d), gives the same quality as Fig. 11(b).

V. EXPERIMENTAL RESULTS

In order to demonstrate that the proposed DIS system accurately stabilize the sequence of images, two real-time sampled indoor images are used. Fig. 12(a) and (b) shows two sampled images with four binary edge MEAs, which are obtained by the edge detection unit. Fig. 13(a) shows the difference image between Fig. 12(a) and (b). Fig. 13(b) shows the difference image between Fig. 12(a) zoomed by 1.25 times and Fig. 12(b) shifted by (3,-4) and then zoomed by 1.25 times. As shown in Fig. 13(b), the image frame motion is significantly reduced.

VI. CONCLUSION

A new approach to the DIS system using tri-state edge detection method is proposed. Pre-processing by edge detection greatly reduces the computational load in motion detection unit, and at the same time it helps producing the correct motion vector. After estimating motion vectors in the motion detection unit, a motion decision step is considered to suppress the weight of motion vectors which is affected by undesired effects. In the digital zooming unit, the ABI method is used for efficient hardware implementation. The proposed DIS system can be realized by an LSI with appropriate line and field memories, and it does not need a Micom.

ACKNOWLEDGEMENT

The authors wish to express sincere appreciation for the valuable help given in the course of this research by Mr. Chan Kyu Myung, staff researcher of Research Group 1, and other members of Buchun R & D Center, Semiconductor Business, Samsung Electronics Corp.

REFERENCES

[1] J. K. Paik and A. K. Katsaggelos, "Edge Detection Using a Neural Network," *Proc. 1990 Int. Conf. Acoust., Speech, Signal Processing*, pp. 2145-2148, Albuquerque, NM, April 1990.

[2] M. Oshimura, et. al., "VHS Camcorder with Electronic Image Stabilizer," *IEEE Trans. Consumer Elec.*, vol. 35, no. 4, pp. 749-758, November 1989.

[3] K. Uomori, et. al., "Automatic Image Stabilizing System by Full-Digital Signal Processing," *IEEE Trans, Consumer Elec.*, vol. 36, no. 3, pp. 510-519, August 1990.

[4] T. Kinugasa, et. al., "Electronic Image Stabilization for Video Camera Use," *IEEE Trans. Consumer Elec.*, vol. 36, no. 3, pp. 520-525, August 1990.

[5] R. P. Lippmann, "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine*, pp. 4-22, April 1987.

[6] B. Widrow, R. G. Winter, and R. A. Baxter, "Layered Neural Nets for Pattern Recognition," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, no. 7, pp. 1109-1118, July 1988.

[7] D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, 1982.

[8] J. K. Paik, *Image Restoration and Edge Detection Using Neural Networks*, Ph. D. Dissertation, Northwestern University, Dept. of EECS, June 1990.

[9] Y. T. Tse and R. L. Baker, "Global Zoom/Pan Estimation and Compensation for Video Compression," *Proc. 1991 Int. Conf. Acoust., Speech, Signal Processing*, vol. M9.4, Toronto, Ontario, Canada, May 1991.

[10] R. J. Schalkoff, *Digital Image Processing and Computer Vision*, John Wiley and Sons, 1989.

[11] A. N. Netravali and B. G. Haskell, *Digital Pictures-Representation and Compression*, Plenum Press, 1988.

[12] V. R. Algazi, G. E. Ford, and R. Potharlanka, "Directional Interpolation of Images Based on Visual Properties and Rank Order Filtering," *Proc. 1991 Int. Conf. Acoust., Speech, Signal Processing*, vol. M12.25, Toronto, Ontario, Canada, May 1991.

Figure 1. Block diagram of a digital video camera with a digital image stabilization (DIS) system and a field memory (FM).

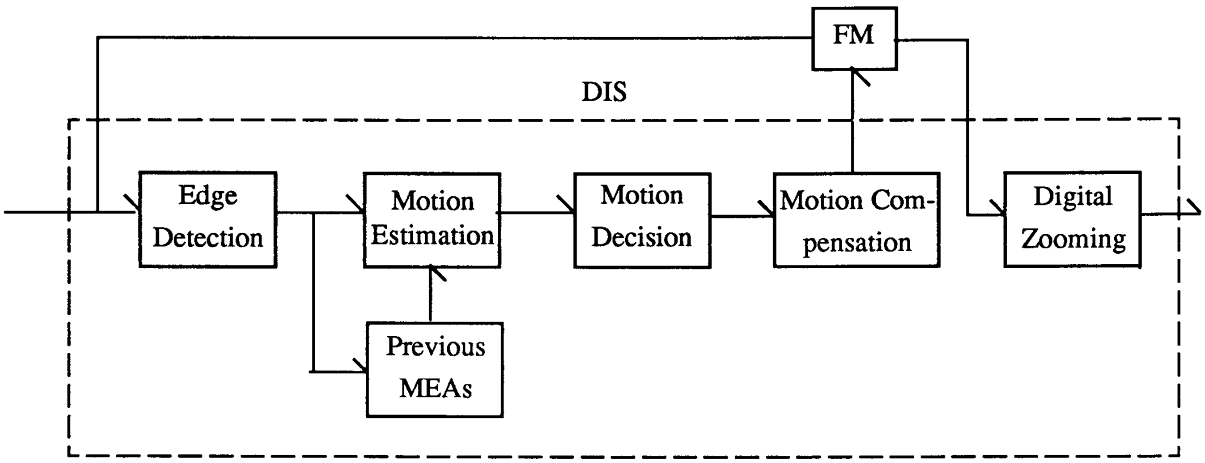


Figure 2. Block diagram of a digital image stabilization system.

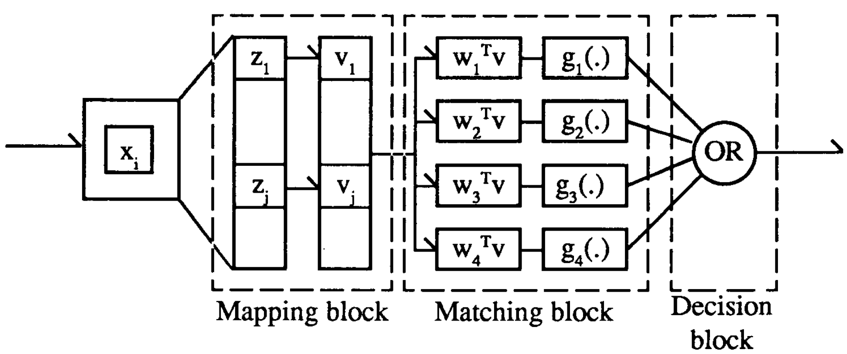


Figure 3. Block diagram of the tri-state ADALINES edge detector.

Figure 4. (a) Original lady image, (b) the edge image of (a) by the proposed edge detection method, (c) Noisy lady image with additive white Gaussian noise (SNR=10dB), (d) the edge image of (c) by the proposed edge detection method, in raster scanning order.

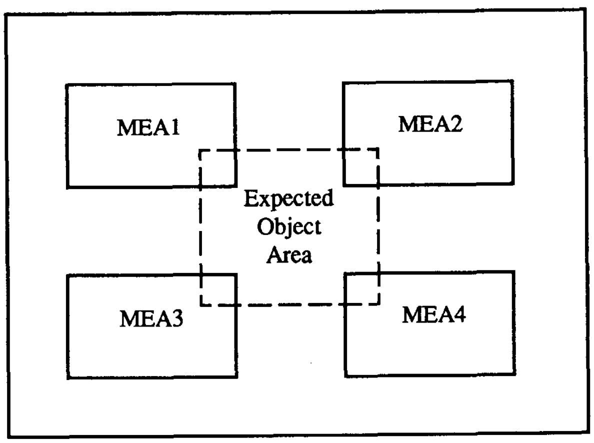


Figure 5. Motion Estimation Areas (MEAs).

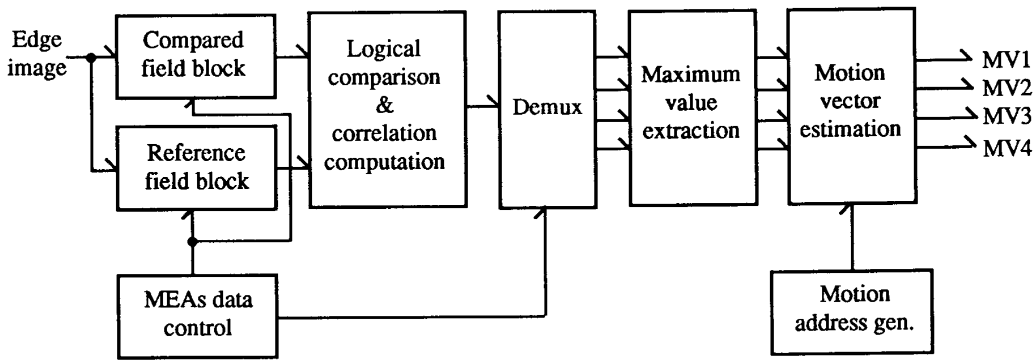


Figure 6. Block diagram of the motion estimation part

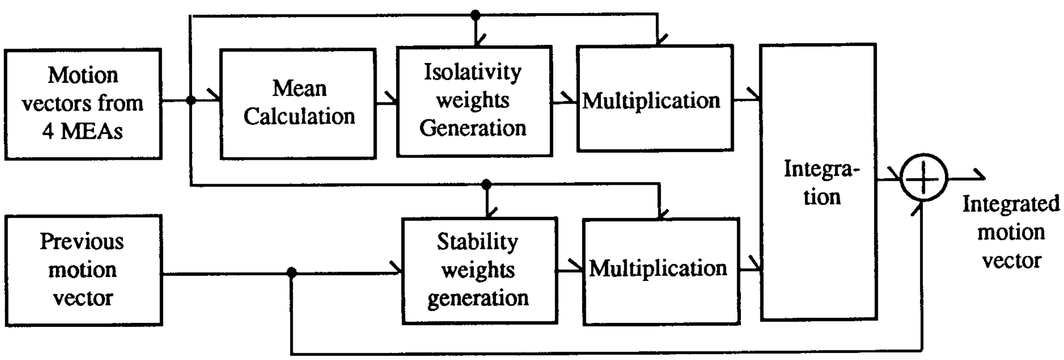


Figure 7. Block diagram of the motion decision part

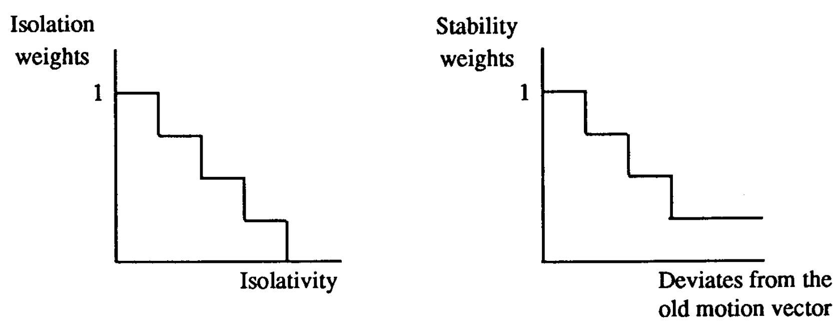


Figure 8. Isolation and stability weights

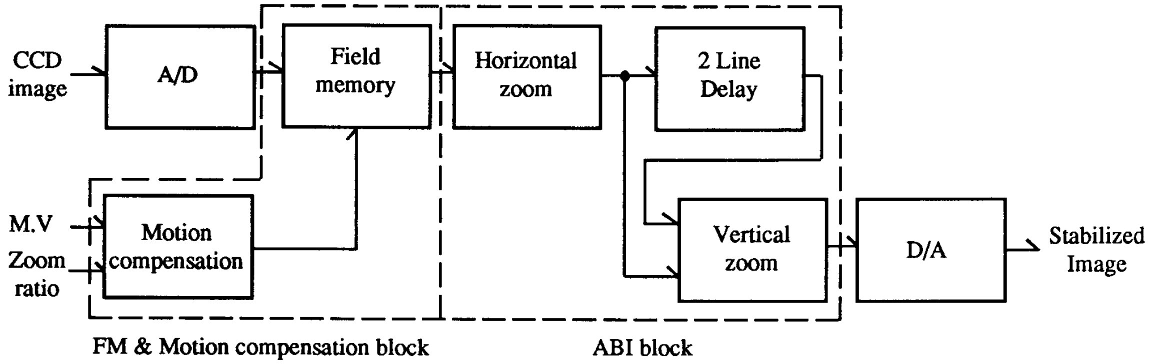


Figure 9. Block diagram of motion compensation and ABI zoom

Original signal line	s_{11}		s_{21}		s_{12}		s_{22}		s_{13}		s_{23}		s_{14}		s_{24}
Zoomed s1 signal line	s_{11}'			s_{12}'			s_{13}'			s_{14}'			s_{15}'		
Zoomed s2 signal line			s_{21}'			s_{22}'			s_{23}'			s_{24}'			s_{25}'

Figure 10. Horizontal zooming of s1 and s2 signals for 1.25 times zooming

Figure 11. (a) Original cameraman image, (b) 4 times zoomed image by the original bilinear interpolation method, (c) 4 times zoomed image by the NNI method, and (d) 4 times zoomed image by the proposed ABI method, in raster scanning order.

Figure 12. (a) The first indoor image with four binary edge MEAs and (b) the second indoor image with four binary edge MEAs, respectively.

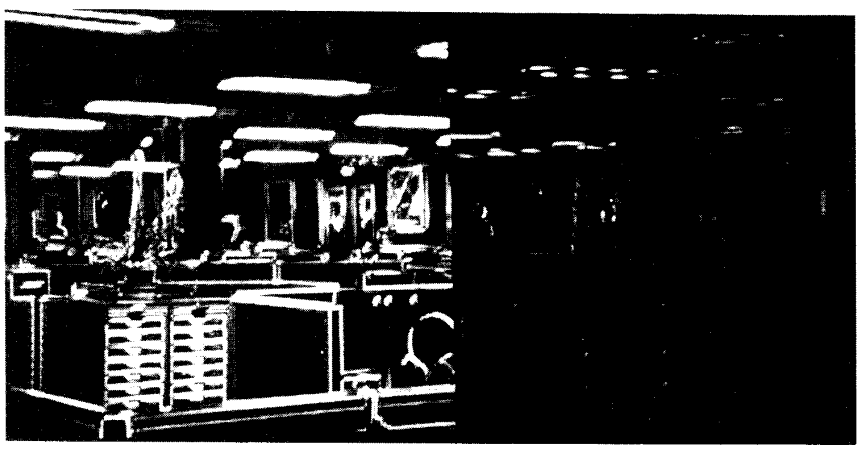
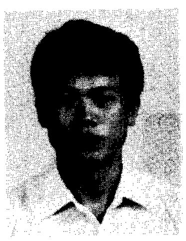


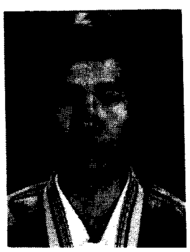
Figure 13. (a) The difference image between Fig. 12(a) and (b), and (b) the difference image between Fig. 12(a) 1.25 times zoomed and Fig. 12(b) shifted by (3,-4) and then 1.25 times zoomed, respectively.

Joon Ki Paik, was born in Seoul, Korea, on July 10, 1960. He received the B. S. degree in control and instrumentation engineering, Seoul National University, in 1984, and the M. S. degree and the Ph. D. degrees in Electrical Engineering in 1987 and 1990, respectively, both from Northwestern University, Evanston, IL.

From January to October 1984, he worked as a researcher in the central research laboratory in Goldstar Instrumentaiton Co. in Korea. From 1988 to 1989, he was a Research Assistant at the Department of electrical engineering and computer science at Northwestern University. He is currently a senior researcher in Video System Tech. Team, Buchun R&D Center, Semiconductor Business, Samsung Electronics Co., where he leas the project for developing the digital image stabilization system for digital video camera. His research interests include signal and image processing, parallel implementation of large scale optimization problems, and artificial neural network implementations of image processing algorithms.



Yong Chul Park was born in 1963, in Iksan, Korea. He received the B. S. degree in Electronics Engineering from Jeonbuk University, Jeonbuk, Korea, in 1989. He joined Samsung Electronics Co. in 1989. His research interests include image processing for digital TV and digital video camera.



Sung Wook Park was born in 1964, Kyungbuk, Korea. He received the B. S. degree in Electrical Engineering from Kyungbuk University, Kyungbuk, Korea, in 1990. He joined Samsung Electronics Co. in Buchun, Korea, in 1990, where he worked on the development of digital auto focus in video camera. Now he is working on the development of digital zooming for digital video camera.