

Non-linear spatial warping of endoscopic images: an architectural perspective for real time applications

Vijayan K. Asari

Department of Electrical and Computer Engineering, Old Dominion University, 231 Kaufman Hall, Norfolk, VA 23529, USA

Received 14 August 2001; revised 7 January 2002; accepted 1 February 2002

Abstract

Images captured with a typical endoscope show spatial distortion, which necessitates spatial warping for subsequent analysis. In this paper, an efficient architecture for an embedded system for the real-time correction of barrel distortion in endoscopic images is proposed. The spatial warping procedure follows a methodology based on least-squares estimation to correct the non-linear distortion in the endoscopic images. A mathematical model of polynomial mapping is used to map the images from distorted image space onto the warped image space. The model parameters include the expansion polynomial coefficients, distortion centre, and corrected centre. The spatial warping model is applied to several gastrointestinal images. The spatial warping algorithm is mapped onto a linear array of simple processing elements with each element of a particular segment communicating with its nearest neighbours. Currently, a prototype of the VLSI architecture for an image of size 256×192 is being designed and built. The functional simulation results obtained in the warping architecture are encouraging. The VLSI based system will facilitate the use of a dedicated module that could be mounted along with the endoscopy unit. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Endoscopic images; Spatial warping; Polynomial mapping

1. Introduction

Electronic video-endoscopy has become one of the commonly accepted forms of diagnostic and therapeutic procedures due to the advent of miniature CCD cameras and associated microelectronics. Video-endoscopes facilitate observation, documentation, and electrical manipulation of the images of internal structure of the gastrointestinal tract. In these endoscopes, cameras with wide viewing angle lens (fish-eye lens) are used to enhance the imaging capability, which permits capturing of larger field in a single image [1]. However, it has been noted that the images obtained from electronic endoscopes show barrel type spatial distortion due to wide-angle configuration of the camera lens. Barrel distortion introduces non-linear changes in the image due to which image areas near the distortion centre are compressed less while areas farther from the centre are compressed more. Because of this, the outer areas of the image look significantly smaller than their actual size. This inhomogeneous image compression introduces significant errors in the results obtained during feature extraction. Continuous estimation of quantitative parameters such as area and perimeter is of considerable

importance while performing clinical endoscopy. Unless the distortion is corrected, estimation errors could be very large [2,3]. In addition, the distortion causes complications while using token matching techniques for pattern recognition. Spatial warping is also a prerequisite for the camera calibration to obtain extrinsic and intrinsic camera parameters [4,5].

Several researchers have presented various mathematical models of the image distortion and techniques to find the model parameters to complete the spatial warping procedure. Tsai [6] proposed a radial lens distortion model that describes a two-dimensional image correction technique. A prism distortion model was used in Ref. [7] to correct the tangential distortion in an image. Nomura et al. [8] presented a calibration technique for high distortion TV camera lens. But this method requires a precise placement of the calibration chart. Thus, a small shift of the chart prompts considerable errors in spatial warping. Weng et al. [9] has explained radial, decentring and thin prism type of distortions and techniques to model them mathematically. All the earlier models give reasonable results for images obtained from cameras with normal viewing objective lens but these models are not effective for electronic endoscopes which use wide-angle lens camera. Smith et al. [10] gave a formulation in which distortion was assumed to

E-mail address: vasari@odu.edu (V.K. Asari).

be purely radial, and orthogonal Chebyshev polynomials were used to determine the model parameters. Hideaki et al. [11] presented a different method for estimation of the model parameters in which a moment matrix was obtained from a set of image points and distorted grid lines in the image were straightened on the basis of smallest characteristic root of the moment matrix. Another technique of distortion correction based on least-squares estimation to obtain the coefficients of the correction polynomial is proposed by Asari et al. [12]. This approach is simpler and faster than the earlier methods and is independent of the orientation of the calibration chart. Thus, it does not require precise placement of the chart and the placement errors in spatial warping formulation can be avoided. It was observed that the reliance on the implementation of the spatial warping algorithm on a PC platform for real-time applications is not possible due to the unacceptable processing time. Thus, it was decided to map this technique to hardware for achieving both high performance and highly integrated solution. It is envisaged that an efficient VLSI implementation will facilitate the placement of this dedicated hardware module within the endoscopy unit.

Low and medium levels of image processing are characterised by a large amount of data to be processed and by classes of algorithms that present a fine/medium level grain parallelism. These two computational characteristics, together with possible real-time applications, lead to consideration of the definition of dedicated architectures that execute the described processing with high efficiency [13]. Due to the excessive computations associated with image processing, many hardware implementation schemes have been presented in the literature to facilitate high-speed performance. A pipelined architecture for image segmentation by adaptive progressive thresholding has been presented by Asari et al. [14]. The segmentation algorithm is mapped onto a linear pipelined architecture in which the computation is fully overlapped with I/O to perform online segmentation. The implementation of systolic architectures for Hopfield and Hamming neural networks are presented in Ref. [15]. A methodology for the design of modular and optimised architectural blocks for the generation of local windows of pixels is presented by Antola et al. [16]. A VLSI implementation of a focal plane image processor for the realisation of the near-sensor image processing concept has been proposed by Eklund et al. [17]. A methodology for creating dense integrated processing element array to build pixel-parallel image processing hardware for micro-computer systems is described by Sodini et al. [18]. All these techniques are aiming towards the realisation of various compute-intensive algorithms to dedicated architectures.

Design of a dedicated architecture for non-linear spatial warping of endoscopic images is proposed in this paper. The design procedure partitions the entire spatial warping process into several functional modules. The special purpose VLSI architecture uses an efficient mapping strategy that reduces the processing time as well as the commu-

nication time between the processing modules. Each processing cell in the pipelined array is simple and requires minimal control logic and the entire architecture could be realised in a single VLSI chip.

2. The non-linear spatial warping algorithm

The spatial warping technique presented in this paper is based on L_2 -norm approximation, which assumes that the distortion is radial about the distortion centre [12]. Although non-linear magnification of the distorted endoscopic image in two dimensions is needed to correct the barrel distortion, the assumption precludes the loss of generality, as a typical endoscope lens is circularly symmetric within narrow precision limits.

Let the distorted and warped image spaces (WISs) be represented by (U', V') and (U, V) respectively, and the distortion centre and the corrected centre by (u'_c, v'_c) and (u_c, v_c) . The distortion centre (u'_c, v'_c) is a point in the distorted image space (DIS) such that the straight lines in the object space passing through it remain straight in the image space. The corrected centre (u_c, v_c) is a point in the WIS about which the expansion of distorted image gives final warped image. In DIS, magnitude ρ' of a vector $\mathbf{P}'$ from the distortion centre to any pixel location (u', v') is given by:

$$\rho' = \sqrt{(u' - u'_c)^2 + (v' - v'_c)^2} \quad (1a)$$

The angle θ' made by the radial vector $\mathbf{P}'$ from the horizontal U' -axis is given by,

$$\theta' = \arctan\left(\frac{v' - v'_c}{u' - u'_c}\right) \quad (1b)$$

Let the same pixel be assigned to a new location (u, v) in WIS; then the magnitude ρ and argument θ of the corresponding vector $\mathbf{P}$ drawn from the corrected centre to the new pixel location are:

$$\rho = \sqrt{(u - u_c)^2 + (v - v_c)^2} \quad (2a)$$

and

$$\theta = \arctan\left(\frac{v - v_c}{u - u_c}\right) \quad (2b)$$

The objective of the mathematical model is to obtain a relation between the vectors $\mathbf{P}'$ and $\mathbf{P}$. An expansion polynomial of degree N is defined to relate the magnitudes of the two vectors in the distorted and warped images as:

$$\rho = \sum_{n=1}^N a_n (\rho')^n \quad (3)$$

where a_n s are the expansion coefficients. As the distortion has been assumed to be purely radial, there will be no change in the arguments of the corresponding vectors $\mathbf{P}'$

and $\mathbf{P}$, i.e. $\theta' = \theta$. After obtaining the magnitude of the new vector, the new pixel location in the WIS can be calculated as:

$$u = u_c + \rho \cos \theta' \quad (4a)$$

$$v = v_c + \rho \sin \theta' \quad (4b)$$

To map each pixel from the DIS onto the WIS, there are $N + 4$ parameters viz. N expansion coefficients (a_n s), distortion centre (u'_c, v'_c) and the corrected centre (u_c, v_c).

2.1. Estimation of expansion coefficients

The expansion coefficients are estimated on the basis of degree of straightness of the points, which lie on a straight line before imaging. These are estimated in the DIS by straightening the grid lines of a distorted grid image. For this purpose an experimental grid is used which contains test dots arranged in horizontal and vertical grid lines. Let P_{ij} denote the centre of a test dot lying in the i th row and j th column of the grid with its co-ordinates at (x_{ij}, y_{ij}) . Let there be L columns of test dots in the grid image with k_j dot centres in the j th column. A set S_j consisting of test dot centres of the j th column is defined as:

$$S_j = \{P_{1j}, P_{2j}, \dots, P_{k_j j}\} \quad \text{for } j = 1, 2, \dots, L \quad (5)$$

To obtain a best fit polynomial curve for each set S_j , a polynomial of degree M is defined as:

$$R_j(x) = \sum_{\alpha=0}^M b_{\alpha j} x^\alpha \quad (6)$$

To estimate the coefficients $b_{\alpha j}$ s, least squares estimation is used which provides sufficient emphasis on all those points which are far from the approximation, without allowing them to dominate. The unknowns, $b_{\alpha j}$ s, are chosen to minimise the function F_j which is defined as:

$$F_j = \left(\sum_{i=1}^{k_j} \left(y_{ij} - \sum_{\alpha=0}^M b_{\alpha j} x_{ij}^\alpha \right)^2 \right)^{1/2} \quad (7)$$

Hence, $b_{\alpha j}$ s can be calculated from

$$\frac{\partial F_j}{\partial b_{\alpha j}} = 0 \quad \text{for } \alpha = 0, 1, \dots, M \quad (8)$$

For every set S_j , from Eq. (6) $M + 1$ simultaneous equations are obtained, which can be represented in a matrix form as

$$\mathbf{H}\mathbf{b} = \mathbf{z} \quad (9)$$

where,

$$\mathbf{H} = [h_{ls}]_{(M+1) \times (M+1)}, \quad h_{ls} = \sum_{i=1}^{k_j} x_{ij}^{l+s}$$

for $l, s = 0, 1, \dots, M$

$$\mathbf{z} = [z_0, z_1, \dots, z_M]^t, \quad z_s = \sum_{i=1}^{k_j} y_{ij} x_{ij}^s$$

$$\mathbf{b} = [b_{0j}, b_{1j}, \dots, b_{Mj}]^t$$

To find the best linear fit for the set of k_j points in j th column of the test dots, a first degree polynomial is obtained from Eq. (6). Hence, two optimum polynomial coefficients are computed using Eq. (9). A normalised error function e_j is defined as the normalised sum of magnitudes of the perpendiculars drawn from each of the k_j points on the best linear fit of j th column as:

$$e_j = \frac{1}{k_j} \sum_{i=1}^{k_j} \left| \frac{b_{1j} x_{ij} - y_{ij} + b_{0j}}{(1 + b_{1j}^2)^{1/2}} \right| \quad (10)$$

The total error for the whole grid image is obtained by:

$$E = \sum_{j=1}^L e_j \quad (11)$$

In the ideal condition when there is no distortion in the image, the total error E is zero since all grid lines will be imaged as straight lines. But due to image distortion, E has a positive value, which decreases monotonously as the distortion reduces. The main objective of the mathematical model is to find the expansion coefficients a_n s to minimise the total error E . Minimisation of E is carried out by an iterative procedure in which the new co-ordinates of the test dot centres in the distorted grid image are calculated by using a new set of expansion coefficients. The recursive relationship to find the new set of expansion coefficients is derived from the line search method of guaranteed convergence [12]. This method is based on a 'globalisation strategy' to select the new coefficients. The search direction in this strategy is different from that derived from the Taylor series as the Taylor series provides local approximation of a function. The global convergence ensures convergence of a series from any starting point to a stationary point. Line search methods are widely used for the purpose of global convergence. On the basis of this method, the expansion coefficients can be obtained by using the following recursive relationship:

$$a_n(\Delta + 1) = a_n(\Delta) - \alpha n^\beta E(\Delta) \frac{1}{\left(\frac{\partial E(\Delta)}{\partial a_n} \right)} \quad (12)$$

for $n = 1, \dots, N$

where α is the convergence rate parameter, β is the expansion index, and $\partial E / \partial a_n$ is the error gradient. Here α is chosen to ensure that for every $(\Delta + 1)$ th iteration, $E(\Delta + 1) < E(\Delta)$. If α is large, the decrease in total error E predicted by the linear approximation may greatly differ from the actual decrease and the global convergence could be violated. On the other hand if α is too small, the time

taken for convergence may be too long. The value of β controls the overall expansion of the distorted image and facilitates the generation of the weights for each of the N coefficients. The iterative relationship given in Eq. (12) also conforms to the principle of diminishing convergence, which ensures fast convergence at the initial phase of the correction procedure when E is significantly large. The iterative procedure ends when the total error becomes smaller than a pre-specified limit ε , i.e. $E(\Delta) \leq \varepsilon$. In the process of spatial warping, E decreases as the grid lines get straightened and it reaches a minimum value $E_{\min}$ when the most optimally straightened grid is obtained. If the image is expanded further, E starts increasing. Thus, if ε is chosen less than $E_{\min}$, the diverging trend in total error is avoided by stopping the iterations further when $E(\Delta + 1) \geq E(\Delta)$.

2.2. Estimation of back mapping polynomial coefficients

Once the expansion coefficients are computed, all the pixels contained in the DIS are mapped onto the WIS. It can be observed that a number of pixel locations are left vacant in the WIS due to the inhomogeneous expansion of the distorted image. To obtain the correct intensity information of these vacant pixels, a back mapping polynomial is derived which maps every pixel from the WIS onto the DIS. This polynomial is defined in a way similar to the expansion polynomial of Eq. (3) and the coefficients are calculated by using non-linear regression analysis employing least squares for a finite number of points in the distorted image [12]. For every pixel in the warped image, the corresponding location in the distorted image is obtained and the information contained in that pixel location is assigned to the warped image pixel. In case, the pixel positions calculated using the back mapping polynomial are non-integers, a linear interpolation on the surrounding pixels is used to get the approximate pixel information.

2.3. Estimation of distortion centre and corrected centre

A reasonably correct estimation of the distortion centre is essential for effective determination of the expansion coefficients. The distortion centre is a fixed point for a particular camera, and once calculated, can be used for all the images obtained from that camera. It can be noted that the curvature of columns decreases as it approaches the centre of the image and then it increases and the sign of the curvature changes. The same observation holds true for the rows also. Thus, the lines, which remain straight after imaging, must lie between the adjacent rows and columns of opposite curvatures. The intersection of such straight lines gives the distortion centre of the image. In the proposed method of distortion centre estimation two polynomials, each of degree γ , passing through a set of grid dots of the adjacent rows of opposite curvatures are defined as:

$$q_k^{(r)}(x) = \sum_{l=0}^{\gamma} c_{kl}^{(r)} x^l \quad \text{for } k = 1, 2 \quad (13)$$

Similarly, two other polynomials, $q_k^{(c)}(x)$ for $k = 1$ and 2, represent the two adjacent columns of opposite curvatures. The best-fit polynomial coefficients $c_{kl}^{(r)}$ are obtained using an equation similar to Eq. (9). The curvature $\kappa_k^{(r)}$ of k th row is computed at the stationary point on a curve, $(x_k^{(r)}, y_k^{(r)})$ as:

$$\kappa_k^{(r)} = \frac{\left. \frac{d^2 q_k^{(r)}(x)}{dx^2} \right|_{(x_k^{(r)}, y_k^{(r)})}}{\left\{ 1 + \left(\left. \frac{dq_k^{(r)}(x)}{dx} \right|_{(x_k^{(r)}, y_k^{(r)})} \right)^2 \right\}^{3/2}} \quad \text{for } k = 1, 2 \quad (14)$$

Similar polynomials $q_k^{(c)}(x)$ for $k = 1, 2$ are also defined for the adjacent columns, and the column curvatures $\kappa_k^{(c)}$ for the k th column at the stationary point $(x_k^{(c)}, y_k^{(c)})$ are obtained. The distortion centre (u'_c, v'_c) is estimated by interpolating the four curvatures $\kappa_1^{(r)}, \kappa_2^{(r)}, \kappa_1^{(c)}$ and $\kappa_2^{(c)}$ as:

$$u'_c = \frac{\kappa_1^{(c)} x_2^{(c)} + \kappa_2^{(c)} x_1^{(c)}}{\kappa_1^{(c)} + \kappa_2^{(c)}}, \quad v'_c = \frac{\kappa_1^{(r)} y_2^{(r)} + \kappa_2^{(r)} y_1^{(r)}}{\kappa_1^{(r)} + \kappa_2^{(r)}} \quad (15)$$

The warped image centre is needed for back mapping as all the vectors $\mathbf{P}$ in the warped image are obtained with respect to this centre. To find this centre, a pixel location is computed in the distorted image, which after spatial warping corresponds to the warped image centre. This pixel location is found based on the criterion that in the warped image, pixel distances between the dot centres should be the same for all the grid lines in the horizontal and vertical directions. The warped image centre is estimated by applying the expansion polynomial to this pixel location in the distorted image, which is obtained by iteratively minimising the variation in distances between the test dot centres in the warped image.

3. Spatial warping architecture

The design of the spatial warping architecture assumes the presence of $2N + 4$ warping parameters viz. N expansion coefficients (a_n s), N back mapping coefficients (b_n s), distortion centre (u'_c, v'_c) and corrected centre (u_c, v_c) . The main activity in the spatial warping procedure is back mapping. Different computational steps involved in the back mapping algorithm described in Section 2 have been mapped into suitable architectural formats along with other necessary algorithmic constituents. The warping algorithm is divided into different computational modules according to their functional sequences and hardware feasibility. The entire architecture is designed with simple computing modules to make the design modular and regular. The warped image could be an expanded image. A scaling circuitry is designed to keep the size of the warped image same as the original image. Conversion of the spatial warping module to an enhanced structure to handle a larger image is possible by the addition of similar modules. The main constituents of the spatial warping architecture are the

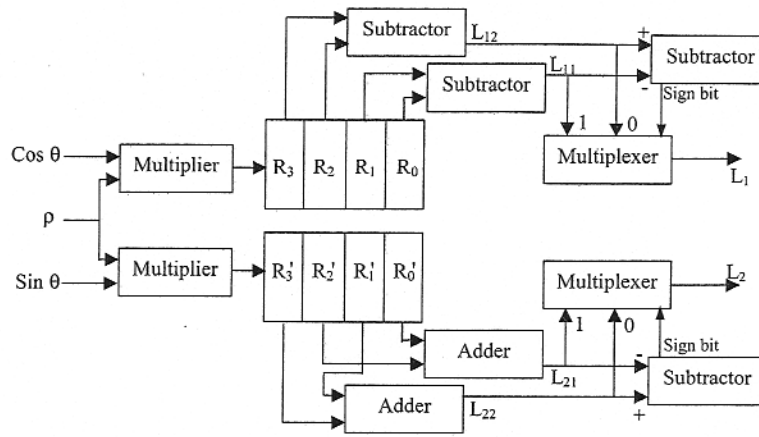


Fig. 2. Architecture for computation of expanded image size.

centre (u_c, v_c) to the pixel location at (l_1, l_2) in WIS and $\cos(l_1, l_2)$ represents the corresponding $\cos \theta$ value. Similarly, register array consisting of R'_0, R'_1, R'_2 , and R'_3 holds the products of $\rho(0,0)$ and $\sin(0,0)$, $\rho(L'_1,0)$ and $\sin(L'_1,0)$, $\rho(0,L'_2)$ and $\sin(0,L'_2)$, and $\rho(L'_1,L'_2)$ and $\sin(L'_1,L'_2)$ respectively. The maximum value of L_{11} and L_{12} is selected by the subtractor–multiplexer configuration, where the sign bit of the subtractor is used as the selection signal for the multiplexer and the width of the expanded image L_1 is obtained as its selected output. Similarly L_2 is obtained from the second multiplexer.

3.2. Module for back mapping

Once the expanded image size is obtained by forward mapping, the pixel locations in this image should be filled with the corresponding pixel values extracted from the distorted image by back mapping. The radius ρ of each pixel location from the corrected centre can be obtained by Eq. (2a) where u takes all values from 0 to L_1 and v takes all values from 0 to L_2 . The cosine and sine values of the corresponding angles are obtained directly as:

$$\cos \theta = \frac{u - u_c}{\rho} \quad (19a)$$

$$\sin \theta = \frac{v - v_c}{\rho} \quad (19b)$$

For the conversion of Cartesian co-ordinate values to the corresponding polar co-ordinate values the same architecture shown in Fig. 1(a) can be used with (u, v) as the input

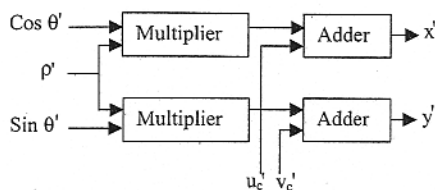


Fig. 3. Computation of the pixel position in DIS.

and (u_c, v_c) as the reference. The radius ρ' of the pixel location in DIS corresponding to each ρ in WIS is obtained by using the back mapping coefficients b_i s as the reference inputs to Fig. 1(b). The polar co-ordinate values of the pixel location obtained in the DIS are transformed into the corresponding Cartesian co-ordinate values by the implementation of the following equations.

$$x' = u'_c + \rho' \cos \theta' \quad (20a)$$

$$y' = v'_c + \rho' \sin \theta' \quad (20b)$$

The hardware schematic for obtaining (x', y') is shown in Fig. 3, where the distortion centre (u'_c, v'_c) is kept as the reference inputs. This pixel location obtained in the DIS can be a fractional value too. The pixel value corresponding to the pixel location (x', y') is computed by linear interpolation of the neighbouring pixel values in DIS.

3.3. Module for linear interpolation

The integer parts of the x' and y' are the x and y co-ordinates of the first pixel location for linear interpolation and they are denoted as A and B , respectively, as represented in Eq. (21a), where the symbol $[p]$ represents the lower bound integer of the real number p .

$$A = [x'], \quad B = [y'] \quad (21a)$$

The fractional parts of the co-ordinate values are represented as A' and B'

$$A' = x' - A, \quad B' = y' - B \quad (21b)$$

The four neighbouring pixel locations are (A, B) , $(A + 1, B)$,

A	A+1
B	B
A	A+1
B+1	B+1

Fig. 4. Pixel locations for linear interpolation.

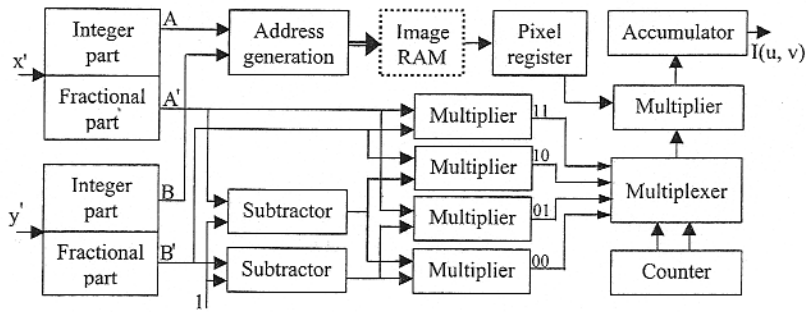


Fig. 5. Pixel value computation using linear interpolation.

$(A, B + 1)$ and $(A + 1, B + 1)$ as shown in Fig. 4. The linear interpolation of the four pixels is performed as:

$$I'_1 = I'(A, B) \times (1 - A') \times (1 - B') \quad (22a)$$

$$I'_2 = I'(A + 1, B) \times A' \times (1 - B') \quad (22b)$$

$$I'_3 = I'(A, B + 1) \times (1 - A') \times B' \quad (22c)$$

$$I'_4 = I'(A + 1, B + 1) \times A' \times B' \quad (22d)$$

$$I(u, v) = \sum_{j=1}^4 I'_j \quad (23)$$

The hardware architecture to obtain the value of $I(u, v)$ by linear interpolation is shown in Fig. 5. The addresses corresponding to the four pixels (A, B) , $(A + 1, B)$, $(A, B + 1)$ and $(A + 1, B + 1)$ in the image RAM are generated and sent sequentially to access the four pixel intensities $I'(A, B)$, $I'(A + 1, B)$, $I'(A, B + 1)$ and $I'(A + 1, B + 1)$ from the RAM. The four fractional product terms are generated by the multipliers and they are fed to the final multiplier using a 4×1 multiplexer. The 2-bit selection input of the multiplexer is from a 2-bit counter whose count sequence is synchronised with the RAM reading operation and shifting of the pixel intensity data from the pixel register. The partial products are accumulated and after four cycles,

$I(u, v)$ is available at the accumulator output. $I(u, v)$ is computed for all values of u and v such that $0 \leq u \leq L_1$ and $0 \leq v \leq L_2$.

3.4. Banked memory architecture

Since massive computations are involved in the spatial warping architecture, it is required to perform simultaneous access and manipulations of image data to enable online performance of the embedded warping module for real-time applications. Hence a banked memory architecture is proposed, which allows multiple data accesses at the same time to perform spatial warping at different portions of the image simultaneously. The basic structure of the pipelining is shown in Fig. 6. The original image data is stored in M different RAMs $RM_1, RM_2, \dots, RM_M$ of capacity $(L_1 + 1) \times (L_2 + 1)/M$ each. The first $(L_2 + 1)/M$ rows of the image are stored in RM_1 . The next $(L_2 + 1)/M$ rows of image in RM_2 and so on and the last $(L_2 + 1)/M$ rows of image are in RM_M . The chip-select (CS) signals for the RAMs are created with the upper significant bits of the original address bus by a suitable CS logic circuitry. When one of the last row elements of a RAM forms the first pixel $I'(A, B)$ for performing linear interpolation, the pipeline requires data from the next RAM locations to be accessed as $I'(A, B + 1)$ and $I'(A + 1, B + 1)$. In this situation, simultaneous data access by two pipelines may occur from two different locations of the same RAM chip. Providing dual port memories as shown in Fig. 6 to the RAMs efficiently rectifies this situation.

3.5. Scaling of the expanded image

It is obvious that the warped image size is larger than the original distorted image as the pixel areas farther from the distortion centre get expanded. The size of the warped image varies depending on the expansion coefficients and it need not be a regular size too. Hence, it is impossible to pre-compute the warped image size and fix the output image RAM size. In this case, the regularisation of the warped image size has become a necessity. This led to the design of a scaling circuitry, for bringing down the size of the warped image to that of the original image. Scaling depends on the horizontal expansion ratio ξ_h and the vertical

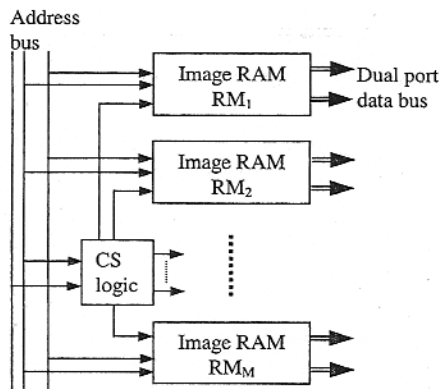
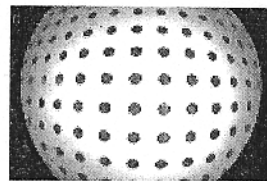
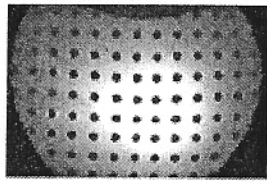


Fig. 6. Schematic for pipelined architecture.



(a)



(b)

Fig. 9. Spatial warping on the experimental grid image: (a) distorted endoscopic image, (b) warped image after back mapping and scaling.

polynomial of order more than 3. The order of expansion polynomial as 4 was chosen. The four expansion coefficients were obtained using the algorithm described in Section 2. The expansion coefficients were computed from the iterative relation given in Eq. (12) for different values of the convergence rate parameter α . The value of the expansion index β was found to be optimum at 2.3, which decides the weight of the n th coefficient. It can be observed that as α increases, convergence becomes faster. As α decreases, $E_{\min}$ reduces while the number of iterations required for convergence increases. But $E_{\min}$ does not decrease significantly after a particular value of α , α_{opt} , though the number of

iterations required for convergence increases exponentially. Hence, to restrict the computation time to make it suitable for online camera calibration, the value of α was chosen as a trade off between the accuracy and the computation time. In the present experimental set up, α_{opt} was chosen as 0.005. For the set of expansion coefficients obtained, the relationship between the distorted radius and the corresponding corrected radius is shown in Fig. 8. It can be observed that the outer areas of the distorted image having large radii expand more after spatial warping.

The back mapping coefficients are found by using a finite number of image dots. For back mapping, the pixel location in the distorted image corresponding to the corrected centre was found as (142, 78). Each pixel in the scaled image space corresponds to a set of pixels in the expanded image and its intensity is obtained by averaging those pixel intensities as shown in Eq. (25). The pixel intensities of the expanded image are obtained by back mapping and linear interpolation. The distorted and warped images of the experimental grid are shown in Fig. 9(a) and (b), respectively. It can be seen that the grid is straightened within practical limits. The same expansion polynomial can be used for spatial warping of the images taken by the video-endoscope until the time the camera lens is not changed or relocated from the CCD array. Typical endoscopic images are taken within a range of less than 20 mm during clinical procedures. The same spatial warping polynomial was applied to several test images captured within a viewing range of 20 mm and it was observed that the warped images were acceptable for further analysis. A typical gastrointestinal image is shown in Fig. 10(a) and the corresponding warped image is shown in Fig. 10(b). It can be noted that the original image was corrected and its outer areas expanded considerably due to the spatial warping.

For the construction of the spatial warping architecture for an image of size 256×192 , the image RAM size was fixed as 4KB. Hence, there were 12 such dual port RAMs to store the image data and 12 similar pipelines for the entire architecture. Each functional module transfers data to its nearest neighbour after completing its computation. The reference data obtained by theoretical analysis of the spatial warping algorithm such as expansion coefficients, back mapping coefficients, distortion centre, and corrected centre are fed to the corresponding reference registers in the architecture. In addition to that, the four corner pixel locations, which define the size of the input image, are also fed to the forward mapping module. The forward mapping module comprising the architectures for co-ordinate conversion and polynomial expansion outputs the new radius and angles corresponding to the four corner pixel positions in the DIS. This data is then fed to the expanded size computation unit, which outputs the new values of L_1 and L_2 . The module consisting of the forward mapping



(a)



(b)

Fig. 10. Spatial warping on a typical endoscopic image. (a) distorted endoscopic image; (b) warped image after back mapping and scaling.

unit and the expanded size computation unit of the architecture is a separate module and it is not a part of the pipeline.

The scaling circuitry is an important part of the architecture, which determines the actual pixel positions to be accessed in the back mapping and linear interpolation modules. The architecture of the back mapping module is the same as the architecture of the forward mapping module except the reference data. The scaling, back mapping, and linear interpolation are performed in all pipelines simultaneously by reading image data from the corresponding image RAMs. A separate timing circuitry for the system is built taking care of the timing requirements of the various computing modules in the architecture. The performance evaluation of the proposed hardware using Altera Quartus II version 1.1 for Apex II FPGA with 20 MHz clock frequency confirms that the overall time latency was about 0.17 ms making it suitable for real-time applications even for a larger image size. The entire architecture is being implemented for eventual fabrication using 0.35 μm CMOS technology. It is envisaged that an improved systolic design with increased parallelism in deep sub-micron technology will significantly shorten the computation time.

The proposed hardware scheme is a generalised architecture, which is suitable for any image size and camera. The camera parameters such as the expansion coefficients, back mapping coefficients, distortion centre, and corrected centre are fed into the system externally. The image size and the image data are obtained from the frame grabber. The corrected image is fed back to the display RAM of the frame grabber for real-time visualisation and diagnosis of the corrected gastrointestinal image by the endoscopist on the endoscope screen.

5. Conclusion

A novel method for the online correction of non-linear distortion in endoscopic images has been presented. The expansion coefficients were obtained by using least-squares estimation and were applied to different grid patterns for testing. It was observed that expansion polynomial obtained for a particular endoscope camera lens was capable of correcting the distortion satisfactorily. This procedure is an essential step for the accurate measurement of the regions of interest in the endoscopic images to facilitate quantitative parameter extraction for online decision-making. Further research is in progress to find a relationship between the expansion coefficients and the distance of the camera lens from the object. The spatial warping algorithm is mapped to an appropriate architecture for building an embedded unit for the endoscopy system to facilitate online camera calibration. A pipelined architecture for the implementation of the

spatial warping technique was implemented such that different computational steps in the algorithm by accessing different portions of the image data are carried out simultaneously. Further research work is in progress to explore a reconfigurable architecture by appropriate partitioning of the different computing modules in the system.

References

- [1] H. Kato, J.P. Barron, *Electronic Videoendoscopy*, Harwood, Switzerland, 1993.
- [2] A. Sonnenberg, M. Giger, L. Kern, C. Noll, K. Stuby, K.B. Weber, A.L. Blum, How reliable is determination of ulcer size by endoscopy, *Br. Med. J.* 24 (1979) 1322–1324.
- [3] C. Margulies, B. Krevsky, M.F. Catalano, How accurate are endoscopic estimates of size, *Gastrointest. Endosc.* 40 (1994) 174–177.
- [4] Q. MaoLin, S. De Ma, Parametric and non-parametric approaches for camera calibration: analysis of imaging errors and their compensation, *Proceedings of the Second Asian Conference on Computer Vision*, Singapore, 1995, pp. II205–II209.
- [5] M. Li, L. Jean-Marc, Some aspects of zoom lens camera calibration, *IEEE Trans. Pattern Anal. Mach. Intell.* 18 (11) (1996) 1105–1110.
- [6] R.Y. Tsai, An efficient and accurate camera calibration technique for 3D machine vision, *Proceedings of the IEEE Computer Vision Pattern Recognition*, Miami, FL, 1986, pp. 364–374.
- [7] D.C. Brown, Decentreing distortion of lenses, *Photogramm. Engng Remote Sensing* (1966) 444–462.
- [8] Y. Nomura, M. Sagara, H. Naruse, A. Ide, Simple calibration algorithm for high-distortion-lens camera, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (11) (1992) 1095–1099.
- [9] J. Weng, P. Cohen, M. Herniou, Camera calibration with distortion models and accuracy evaluation, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (10) (1992) 965–980.
- [10] W.E. Smith, N. Vakil, S.A. Maislin, Correction of distortion in endoscopic images, *IEEE Trans. Med. Imaging* 11 (1) (1992) 117–122.
- [11] H. Hideaki, Y. Yagihashi, Y. Miyake, A new method for distortion correction of electronic endoscope images, *IEEE Trans. Med. Imaging* 14 (3) (1995) 548–555.
- [12] K.V. Asari, S. Kumar, D. Radhakrishnan, A new approach for non-linear distortion correction in endoscopic images based on least squares estimation, *IEEE Trans. Med. Imaging* 18 (4) (1999) 345–354.
- [13] S.Y. Kung, *VLSI Array Processors*, Prentice Hall, Englewood Cliffs, NJ, 1988.
- [14] K.V. Asari, T. Srikanthan, S. Kumar, D. Radhakrishnan, A pipelined architecture for image segmentation by adaptive progressive thresholding, *J. Microprocessors Microsyst.* 23 (8–9) (1999) 493–499.
- [15] K.V. Asari, C. Eswaran, Systolic array implementation of artificial neural networks, *J. Microprocessors Microsyst.* 18 (8) (1994) 481–488.
- [16] A. Antola, A. Avai, L. Breveglieri, Modular design methodologies for image processing architectures, *IEEE Trans. VLSI Syst.* 1 (1993) 408–414.
- [17] J.E. Eklund, C. Svensson, A. Astrom, VLSI implementation of a focal plane image processor—a realisation of the near-sensor image processing concept, *IEEE Trans. VLSI Syst.* 4 (1996) 322–335.
- [18] C.G. Sodini, J.C. Gealow, Z.A. Talib, I. Masaki, Integrated memory/logic architecture for image processing, *Proc. IEEE VLSI Des.* (1997) 304–309.



Vijayan K. Asari received the BSc (Eng) degree in Electronics and Communication Engineering from the University of Kerala, MTech and PhD degrees in Electrical Engineering from the Indian Institute of Technology, Madras, India. He has been working as an Assistant Professor in T.K.M. College of Engineering, University of Kerala, India. He was with the National University of Singapore from October 1996 to June 1998 where he performed research on biomedical image processing and neural networks. He was with the School of Computer Engineering at Nanyang Technological University, Singapore from July 1998 to July 2000. He joined the Department of Electrical and Computer Engineering, Old Dominion University, Virginia as an Associate Professor in August 2000. His research activities include design and development of digital system architectures for real-time embedded systems in the areas of image processing and neural networks. Dr Asari is a Senior Member of IEEE.