Combined Independent Component Analysis and Kalman Filter Based Real-Time Digital Video Stabilization

Hassaan S. Qureshi, Syed A. Jabir, Syeda H. Taqdees, and Khawar Khurshid
Department of Electrical Engineering, School of Electrical Engineering and Computer Sciences
National University of Sciences & Technology (NUST) Islamabad, Pakistan
Email: {11mseehsaadat, 11mseeajabir, 11mseestaqdees, khawar.khurshid}@seecs.edu.pk

***Abstract*—We present an online approach for digital video stabilization. This paper builds upon the work presented in where the authors have proposed the use of Independent Component Analysis (ICA) for video stabilization. Their scheme however is intended for offline systems and cannot be used for online systems. We propose a window based approach to combine the ICA based technique with Kalman filter, in order to make the method suitable for online systems. The proposed approach is tested using simulations and the mean square error between the outcome and the ground truth is used as performance measure. The performances of the proposed method, separate ICA and separate Kalman filter application are compared. The performance of the proposed method turns out to be better than separate application of Kalman filtering or ICA alone.**

***Index Terms*—video stabilization, ICA, Kalman filter**

I. Introduction

Over the recent years, the domain of video stabilization has been active area of research. The techniques for video stabilization are becoming popular with the growing advancement in digital imaging devices.

Most of the times, the base of the camera may not be perfectly fixed. The camera mounted on a traffic signal, on a moving vehicle or hand held may be subjected to various vibrations and swings caused by wind, engine vibrations or hand unsteadiness respectively. The effect of these unintentional camera motion results in annoying high frequency motions in the captured image sequence. Furthermore, a part of the motion is intentional when we intentionally move the camera or the camera carrying car turns. Thus the process of video stabilization is to remove the unintentional motion while keeping the intentional motion. These unintentional high frequency vibrations are not only unpleasing for humans but also reduce the performance of other computer vision and compression algorithms like object tracking and segmentation etc.

The process of video stabilization can be divided into three main steps, 1) Motion estimation, 2) Motion taxonomy or Motion separation and 3) Motion compensation.

In the first stage, we estimate the motion of each frame with respect to the previous frame. This is called motion estimation. At this stage, the motion vector is the mixture of the intentional motion and the unwanted high frequency motion. Step 2 involves the separation the intentional motion from the unwanted motion and finally in Step 3, we apply the inverse of the unwanted motion to the video sequence, while keeping the intentional motion. A lot of research has been done on each of the three steps. In this paper, we focus on the step of motion taxonomy which is the heart of Digital Video Stabilization.

Recently, Amanatiadis and Andreadis [1] have proposed a novel technique for motion separation based on Independent Component Analysis. The authors assume that both intentional and unintentional motions are independent of each other and hence they can be separated by ICA. However, they have mentioned that their technique is not suited for online stabilization as the methodology requires all motion vectors prior to the application of ICA.

In this paper, we propose a modification to the proposed algorithm in [1] to make it suitable for real time implementation. We combine Kalman filter [2] with the proposed approach of ICA to achieve this target. Simulation results show that the output intentional motions vector are less jittery than both Kalman based stabilization output and ICA based stabilization output when they are used separately.

The organization of this paper is as follows. In Section II, we present the existing approaches for digital video stabilization. A brief introduction to ICA and Kalman filtering is given in Section III. The methodology proposed in [1] and our proposed modification is presented in Section IV. In Section V we present the simulation results and finally, we conclude the paper in Section VI with some discussion of the future work.

II. Related Work

In this section, we present different procedures that have already been proposed in literature by different authors.

Manuscript received September 13, 2013; revised November 19, 2013.

doi: 10.12720/ijsps.1.2.164-169

In one design proposed by Ratakonda K. [3], a computational efficient technique is developed to estimate the local and global motion vectors, assuming that small variations can be handled as transitional model using projection matching. However, this paper does not handle the moving object problem.

Piva S. *et al.* [4] propose the real time video stabilization for the camera mounted on a locomotive of a high-speed train. They have applied translational model to find the rotational parameters by assuming large peripheral of camera doesn't change maximum correlation value. The authors in [5] apply the similarity transformation to images obtained by the camcorder mounted on a moving vehicle by estimating motion based on global feature extraction such as road lanes. Zhu J. *et al*. [6] estimated the local and global motion vectors by dividing the image into four sub-images and finding the local motion parameters for each one separately. These local vectors are weighted adaptively in order to find the global motion parameters.

The authors in [7] proposed a design to estimate camcorder motions by calculating the histograms of local motions. The position of the highest peak in each local motion histogram is independent to estimation errors, thus providing a robust solution to the interference of moving objects. Yang J. *et al.* [8] proposed the affine model transformation for the motion estimation obtained by the hand held camera. They facilitated the use of affine model by using the feature point detection and particle filters to estimate the global camera motion between successive frames.

The authors in [9] propose stabilization technique using both the similarity and the affine transforms. The approach is to extract the robust weighted feature trajectories from the input video and find a set of transformations to smooth out these trajectories and finally stabilize the video. Battiato *et al*. [10] proposed a method to perform block matching, pre-filtering, memory filtering, robust estimation and then computation of error vector. Battiato *et al.* [11] have also proposed another approach in which the global motion model for the video frame caters translations in x and y axis, a rotation and zoom factor. Least square method is used to avoid motion estimation errors. In [12], Hu et al. use affine model for global motion model and SIFT is applied to estimate camera motion. Shaky motion and intentional motion are separated out by Gaussian filtering.

III. Theoretical Background

In this section we present some basic introduction to the Kalman filter and the ICA.

A. Kalman Filter

In most simple words, Kalman filtering can be simply described as noisy data in, hopefully less noisy data out [13].

The Kalman filtering consists of two sets of computations, the prediction part, and the correction part. In the prediction part, at time index k-1, the Kalman filter first predicts the process state vector for time index k. This is called as apriori estimate. In the correction stage, the Kalman filter correct this apriori estimate after obtaining the measurement at time index k. This is the posteriori estimate, calculated by giving an optimal weightage to the measurement and the prediction using the Kalman gain.

$$\begin{aligned} \hat{x}_k^- &= A\hat{x}_{k-1} + Bu_k \\ P_k^- &= AP_{k-1}A^T + Q \end{aligned} \tag{1}$$

Above equations are known as priori estimates or predictor equations. $\hat{x}_k^-$ and P_k^- are the state and covariance estimates at index $k-1$. A is the state transition matrix and Q is the process noise covariance.

$$\begin{aligned} K_k &= P_k^- H^T (HP_k^- H^T + R)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \\ P_k &= (I - K_k H)P_k^- \end{aligned} \tag{2}$$

These equations are known as posteriori estimate or corrector equations. K_k is the Kalman gain, z_k is the observed output which is used to estimate state at index k. R is the measurement noise covariance.

In this paper, we have used the constant speed model as it is sufficient for most of the cases.

B. Independent Component Analysis

Independent component analysis is a technique to recover the unobservable independent signals from their observable linear combinations by maximizing their non-gaussianity. It is the extension of Principle Component Analysis (PCA) requiring the additional constraint of Independence of sources along with their uncorrelated-ness [14].

ICA is described using the statistical "latent variables" model of [15]. Suppose that we have observed n random variables $x_1, x_2, ..., x_n$ which are the linear combinations of n independent sources $s_1, s_2, ..., s_n$ such that

$$x_j = a_{j1}s_1 + a_{j2}s_2 + ... + a_{jn}s_n \tag{3}$$

Using matrix notation, observation of n x random variables can be expressed as

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{4}$$

where $\mathbf{x}$ and $\mathbf{s}$ are the vectors containing x_j and s_j respectively. $\mathbf{A}$ is called the mixing matrix which consists of the rows of mixing coefficients for each of the observed linear combinations. More precisely, the columns of $\mathbf{A}$ depict the contribution of each source s_i to the linear combinations x_j i.e.

$$\mathbf{x} = \sum_{i=1}^{n} \mathbf{a}_i s_i \tag{5}$$

In ICA, sources s_i must be non-gaussian as well as independent. It analyzes the relationship between

independent sources and their linear combinations by estimating de-mixing matrix $\mathbf{W}$ (the inverse of the unknown matrix $\mathbf{A}$) by maximizing the non-gaussianity of the pdfs of x_j. The key idea is that the sum of non-Gaussian random variables s_i are much more Gaussian then the original ones based upon the central limit theorem of statistical random variables. So for estimating the latent variables s_i, we try to maximize the non-gaussianity and independence of x_j by multiplying it with a rotational matrix $\mathbf{W}$ such that

$$\mathbf{s} = \mathbf{W}\mathbf{x} \tag{6}$$

Non-gaussianity of random variables are measured by calculating kurtosis i.e. fourth-order cumulate. For a random variable with zero mean, the normalized kurtosis is calculated through the following formula.

$$kurt(u) = \frac{\varepsilon\{u^4\}}{(\varepsilon\{u^2\})^2} \tag{7}$$

The basic property of the normalized kurtosis is that, for Gaussian random variables, kurtosis is zero. For most non-Gaussian random variables, kurtosis is nonzero.

There are some limitations associated with ICA. As both the $\mathbf{A}$ and $\mathbf{s}$ in (4) are unknown, we can freely change the order of the terms in the sum in (5), so we cannot determine the exact order of the independent components. Thus the order of the output signal is uncertain. This ambiguity is called as permutation ambiguity.

Another limitation is that the energies of the resulting independent signals may be a scaled version of energy of the actual signals. This is again because both $\mathbf{A}$ and $\mathbf{s}$ are unknown in (4) and any scalar multiplication with $\mathbf{s}$ can be balanced with the corresponding scalar division with $\mathbf{A}$. This is called amplitude ambiguity.

IV. Proposed Method and Modifications

In this section, we first describe the approach presented by Amanatiadis and Andreadis [1] in detail, then in the next subsection, we present our proposed approach to make the technique suitable for online video stabilization.

A. Video Stabilization Using ICA

As mentioned earlier, the authors in [1] have proposed a novel approach of separating the wanted and unwanted motion from the estimated motion vectors. They assume the two motions are independent of each other, and hence they can be separated by Independent Component Analysis. ICA is basically a technique used for separation of individual signal from a mixture of independent signals mostly popular for the cocktail party problem [16]. The authors proposed that the problem of video stabilization can be made analogous to the cocktail party problem. The two recording microphones correspond to different regions of the frame and the recorded sound corresponds to the estimated motion vectors for each region. The regions which include the foreground objects should not be selected for motion vector estimation. The authors also refer to the different techniques to detect whether a region contains foreground objects or not. Thus the wanted and unwanted motion can be separated by the application of ICA.

As mentioned earlier, the ICA has two major limitations, the permutation ambiguity and the amplitude ambiguity. The authors have also suggested a solution to these problems.

The permutation ambiguity can be resolved by assuming that the wanted motion has lower frequency that the unintentional motion and doing the frequency analysis. After this ambiguity has been resolved, we can easily resolve the magnitude ambiguity by projecting the signal back to the observations space. As a result we get the contribution of each constituent signal, the intentional and unintentional motion, in each of the observation. The signal representing the intentional and unintentional motion can then be obtained by averaging their respective contributions in the observation signals.

However, the authors describe that this algorithm cannot be used for online stabilization as this technique requires motion vectors for all video frames prior to the ICA application.

In the following subsection, we present our approach as the solution to this problem where we use Kalman filter in combination with ICA to modify the algorithm in order to make it suitable for online implementation.

B. Proposed Modifications

As described in Section III, the Kalman filter first predicts the state of the process at time index *k*, and then corrects the prediction after obtaining the measurement at time index *k*. The original Kalman filtering for the motion separation stage [17] proposes the application of Kalman filter on the observed motion vector. The motion vector is taken as the input noisy data and the smooth output motion is taken as the separated intentional motion.

In our method, we use the idea of sliding window and propose to compute ICA for each window separately. If the window size is *W*, then at each frame index *k*, we compute the ICA for (*k*-*W*)th frame to *k*th frame. The value of resultant ICA-separated intentional motion at *k* is given as *k*th measurement to the Kalman filter which then computes the actual value of the curve at that point. Effectively, this means that the Kalman filter is now operating on less noisy data, which will result in smoother output.

For the starting frames, when $k<W$, we assume that the motion vector data is our measurement and is passed directly to Kalman filter. Hence the performance will be relatively worse for the frame when $k<W$, but after this, the motion separation performance of this system provides better results than the cases when ICA or Kalman filter is applied separately for the motion estimation stage.

V. Results

The proposed algorithm of combining ICA with Kalman filter has been tested on simulated data and the results are shown and discussed below.

A. Testing Setup

The wanted motion curve is modeled with a signal that is sum of sines and cosines (Fig. 1). The unintentional motion is simulated by zero mean uniform random process (Fig. 2).

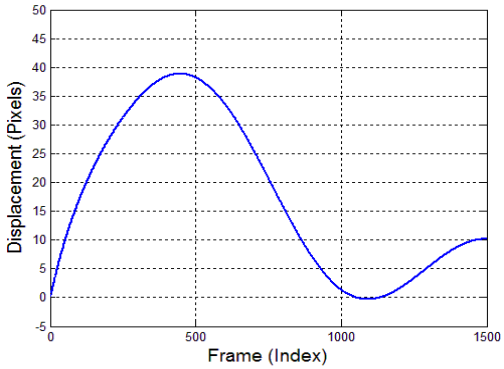

Figure 1. Desired motion curve.

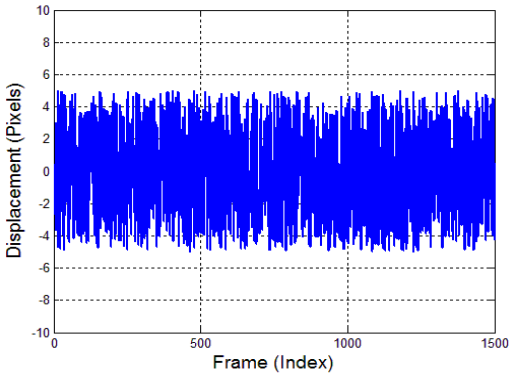

Figure 2. Unintentional motion.

The two signals are mixed using a mixing matrix to get two observations which are then passed to our system (Fig. 3 and Fig. 4). To apply Kalman filter, we need only a single measurement, so we take average of the two observations (Fig. 5). This step is justified as the ICA based algorithm also returns the intentional motion as the average of its contributions in each observed signal. The block size W is set to 50 for the experiments.

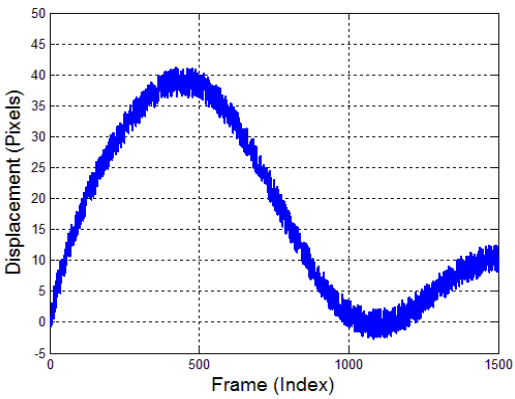

Figure 3. Observation 1.

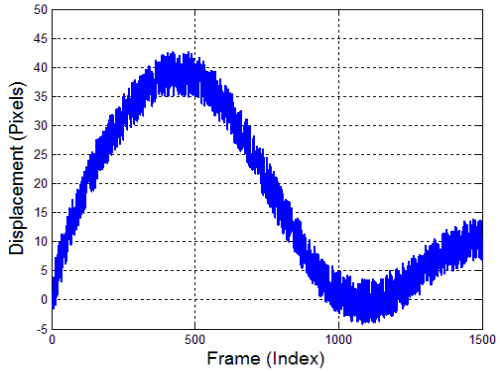

Figure 4. Observation 2.

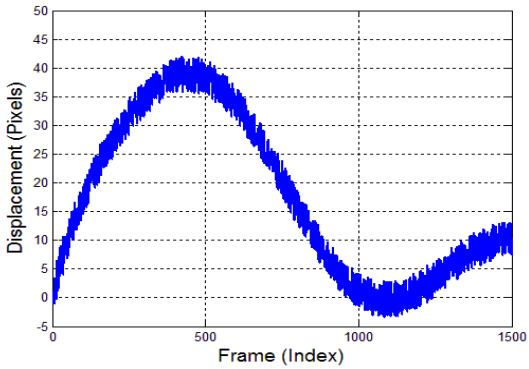

Figure 5. Mean of observation 1 & 2.

B. Discussion

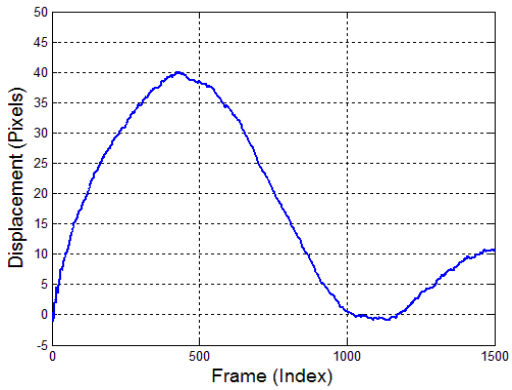

Figure 6. Output of Kalman filter only.

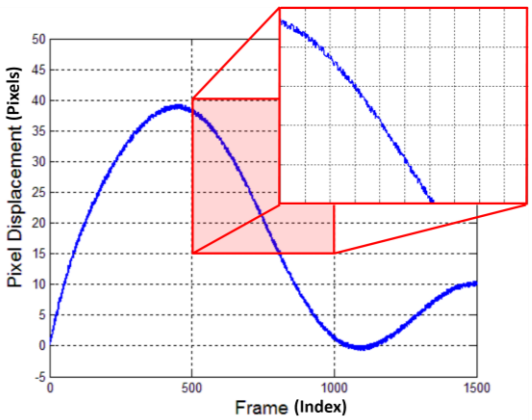

Figure 7. Output of ICA alone.

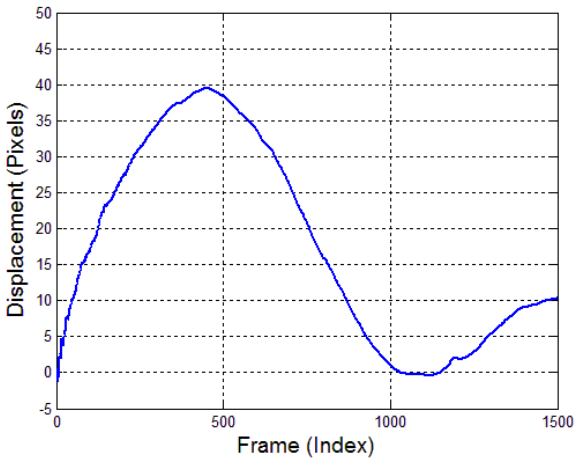

Figure 8. Output for combined Kalman filter and ICA (CICAK).

Fig. 6–Fig. 8 shows the output for the Kalman filter alone, the ICA alone and the proposed combined ICA-Kalman (CICAK).

Clearly, it can be observed the result for the CICAK is the best among the three as far as smoothness is concerned. The output for the CICAK is initially same as the out of Kalman because at those instances, the observed motion vector is given as measurements to the Kalman filter. After *W* samples, the ICA-separated intentional motion is given as measurements to the Kalman filter. Since this data is less noisy, thus the output of the Kalman filter becomes smoother.

We have used first derivative as a smoothness measure. Fig. 9 shows the smoothness curve for Kalman alone, ICA alone, CICAK and GT (ground truth). We can see that after *W* samples CICAK follows GT in a smoother way compared to ICA and Kalman alone. The mean square error between derivative of GT and CICAK is 0.000109.

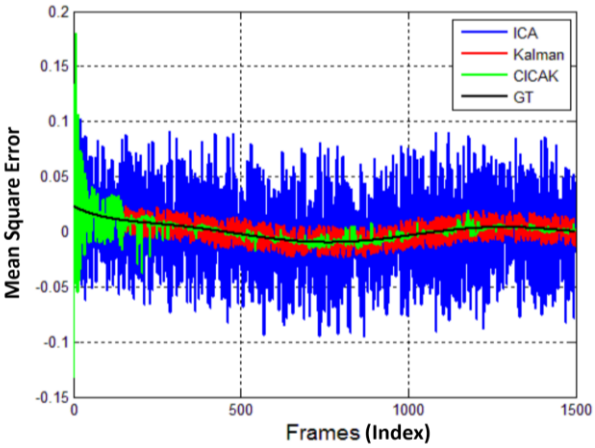

Figure 9. Smoothness curve.

VI. Conclusion and Future Work

In this work, we have presented an online video stabilization method based on ICA and Kalman filter. The work is intended to solve the limitation of non-causality in the work of [1]. Instead of computing ICA for the complete Motion Vector data, we proposed a window based approach, where ICA is computed for the block of data containing the value of motion vector at the present frame and at *W-1* previous frames. The output of this system is passed as measurements to the Kalman filter. The results show that the smoothing performance of this system is better than the case ICA or Kalman filtering is applied separately.

At present we have only tested the approach on simulated data. In future, we plan to test the method on real video. Another thing that can be tested is that we used pure sliding window where each window is shifted by single index at a time. The jump of the window can be increased to save the computational requirements.

References

[1] A. A. Amanatiadis and I. Andreadis, "Digital image stabilization by independent component analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 7, pp. 1755-1763, 2010.

[2] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, no. 82, Series D, pp. 35-45, 1960.

[3] K. Ratakonda, "Real-time digital video stabilization for multi-media applications," in *Proc. IEEE International Symposium on Circuits and Systems*, 1998, pp. 69-72.

[4] S. Piva, M. Zara, G. Gera, and C. Regazzoni, "Color-based video stabilization for real-time on-board object detection on high-speed trains," in *Proc. IEEE Conference on Advanced Video and Signal Based Surveillance*, 2003, pp. 299-304.

[5] Y. Liang, H. Tyan, S. Chang, H. Liao, and S. Chen, "Video stabilization for a camcorder mounted on a moving vehicle," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 6, pp. 1636-1648, 2004.

[6] J. Zhu and B. Guo, "Video stabilization with sub-image phase correlation," *Chinese Optics Letters*, vol. 4, no. 9, pp. 553-555, 2006.

[7] J. Cai and R. Walker, "Robust motion estimation for camcorders mounted in mobile platforms," presented at Digital Image Computing: Techniques and Application 2008, Canberra, Australia, 1-3 December 2008.

[8] J. Yang, D. Schonfeld, and M. Mohamed, "Robust video stabilization based on particle filter tracking of projected camera motion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 7, pp. 945-954, 2009.

[9] K. Lee, Y. Yu, C. Yu, and C. Ouhyoung, "Video stabilization using robust feature trajectories," in *Proc. IEEE International Conf. on Computer Vision*, Japan, September 2009, pp. 1397-1404.

[10] S. Battiato and G. Puglisi, "A robust video stabilization system by adaptive motion vectors filtering," presented at 2008 IEEE International Conference on Multimedia and Expo, Hannover, Germany, June 23-26, 2008.

[11] S. Battiato, G. Gallo, and S. Scellato, "SIFT features tracking for video stabilization," in *Proc. 14th International Conference on Image Analysis and Processing*, 2007, pp. 825-830.

[12] R. Hu, R. Shi, and I. Shen, "Video stabilization using scale-invariant features," in *Proc. 11th International Conf. on Information Visualization*, 2007, pp. 871-877.

[13] G. Bishop and G. Welch. An Introduction to the Kalman Filter. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.163.5348&rep=rep1&type=pdf

[14] J. Karhunen, E. Oja, L. Wang, R. Vigario, and J. Joutsensalo, "A class of neural networks for independent component analysis," *IEEE Transactions on Neural Networks*, vol. 8, no. 3, pp. 486-504, 1997.

[15] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, vol. 13, pp. 411-430, May-June, 2000.
[16] Y. Ainhoren, S. Engelberg, and S. Friedman, "The cocktail party problem," *IEEE Instrum. Meas. Mag.*, vol. 11, no. 3, pp. 44–48, June 2008.
[17] A. Litvin, J. Konrad, and W. C. Karl, "Probabilistic video stabilization using Kalman filtering and mosaicking," in *Proc. SPIE Image and Video Communications and Proc*., 2003, pp. 663-674.

Hassaan S. Qureshi received his B.E. degree in electrical engineering from National University of Sciences & Technology (NUST), Pakistan. He is currently pursuing M.S degree in electrical engineering from School of Electrical Engineering & Computer Science (SEECS), NUST. He is a research assistant at Vision Image and Signal Processing (Vispro) Lab and research interests include image processing, embedded systems and HDL based digital design.

Ms. Hira Taqdees obtained her Bachelors of electrical engineering degree from College of Electrical and Mechanical Engineering (CE&ME), NUST, Pakistan. and is currently pursuing her Masters in electrical engineering from School of Electrical Engineering and Computer Science (SEECS), NUST. Her field of specialization is digital image processing and formal verification of analog circuits. Current and previous research interests ends the paragraph.

Syed A. Jabir received his B.E. and M.S degrees in computer engineering and electrical engineering respectively from National University of Sciences & Technology (NUST), Pakistan. He is currently a lecturer at Center for Advanced Studies in Engineering (CASE). His areas of research interest include signal processing, image processing embedded systems and FPGA based hardware-software co-design.

Dr. Khurshid obtained his doctorate degree from Michigan State University in Biomedical Imaging Systems and currently he is heading the Institute of Applied Electronics and Computers at National University of Sciences and Technology (NUST), Pakistan. He is also the director of Cypress-NUST Research Center focusing on solutions development on PSoC. Dr. Khurshid specializes in Medical Imaging, Computer Vision, Pattern Recognition, Image and Signal Processing with research interests in the areas of image segmentation, registration, video encoding, multimedia streaming and 3D video display systems.