

IMAGE STABILIZATION BY KALMAN FILTERING USING A CONSTANT VELOCITY CAMERA MODEL WITH ADAPTIVE PROCESS NOISE

Eylem Yaman

e-mail: eylem@kou.edu.tr

Sarp Ertürk

e-mail: sertur@kou.edu.tr

Kocaeli University, Faculty of Engineering, Department of Electronics & Telecommunications Engineering, 41100, Izmit, Kocaeli, Turkey

Key words: Image stabilization, Kalman filtering, Frame position smoothing

ABSTRACT

An image sequence stabilization system based on adaptive Kalman filtered smoothing of global camera motion is presented in this paper. The global camera displacement is modelled in the form of a constant velocity motion model, that is applied to the Kalman filter to ensure smooth global displacements. The process noise variance of the Kalman filter is varied adaptively according to former correction vectors, to prevent the correction vector from exceeding the permitted frame shift limit, while achieving successful stabilization.

I. INTRODUCTION

Image sequence stabilization (ISS) systems are used to remove undesired image fluctuations (jitter) from a sequence so as to enhance visual quality [1-16]. Instability of the camera operator or the camera platform may result in the image sequence being subject to involuntary global motion effects, and by removing such undesired global motion components a compensated image sequence displaying only requisite global motions can be obtained. The application field ranges from state-of-the-art camcorders to forthcoming third generation (3G) wireless video communication systems.

Image stabilization systems consist of two parts: the motion estimation part and the motion correction part. The motion estimation part resolves global motions contained in an image sequence, and the motion correction part compensates for the jitter components. Primitive image stabilization systems used to acquire the vibration statistics via mechanical sensors [1,2], and perform stabilization by changing the refraction angle of the optical lens [3]. Current image stabilization systems are now fully digital, motion estimation and correction is performed by digital signal processing [4-16], enabling robust, inexpensive and comparably lighter systems.

This work was supported by The Scientific and Technical Research Council of Turkey under Contract EEEAG-101E006.

Image sequence stabilization systems can compensate for translational, rotational or scale inconsistencies. Visual quality is most severely degraded by translational jitter, and therefore most stabilization systems have focused on translational jitter only [4-11], commonly referred to as two dimensional (2-D) stabilization systems. Stabilization systems that compensate for rotation and/or scale variations with preserved or compensated translational jitter are referred to as three dimensional (3-D) stabilization systems [12-14].

Accurate and fast estimation of global interframe motion, has been the common subject of research for image sequence stabilization systems. Estimation of global motion by block-matching of subimages has been proposed in [4], and the technique has been extended to bit-plane matching in [5] and gray-coded bit-plane matching in [6]. Global motion estimation based on edge pattern matching has been demonstrated in [7]. Multiresolution approaches for global motion estimation making use of affine motion parameters between levels of the Laplacian pyramid, mosaic-based registration, and feature-based matching have been presented in [8], [9] and [10] respectively. A robust utilisation of phase correlation based motion estimation has been presented in [11]. Tracking of visual cues with 3-D motion parameters estimated by Kalman filtering for the compensation of image rotations has been demonstrated in [12]. An image flow algorithm has been used to estimate global motion in [13]. The estimation of 3-D motion parameters from two dimensional affine motion has been presented in [14].

The motion correction part of stabilization systems has received less attention. Several stabilization systems have been implemented to remove all of the estimated global motion [8-10], without resolving the jitter component from intentional global camera motion. Three dimensional stabilization systems are mainly aimed at removing [12-14] or smoothing [15] the rotational motion component. Motion vector integration (MVI) [4,5,6,16] and frame position smoothing (FPS) [11] have been presented to

stabilize translational jitter, with the aim to preserve deliberate camera movements. While MVI is implemented to operate in real-time, it has the drawback of requiring a compromise between stabilization performance and intentional movement preservation. Frame position smoothing based on DFT filtering as proposed in [11], on the other hand accomplishes an optimal balance between stabilization performance and deliberate global camera preservation capabilities, however is only suited for off-line applications due to the Fourier domain implementation. Kalman filtered frame position smoothing is proposed to provide the superior stabilization performance of FPS at real-time.

This paper presents an image sequence stabilization system based on adaptive Kalman filtering of absolute frame positions [17]. The camera global motion is modelled using a constant velocity motion model, and Kalman filtering is employed to ensure smooth changes in frame positions. The process noise variance of the Kalman filtering process is adaptively changed according to the magnitude of previous correction vectors (similar to adaptively changing the damping coefficient of MVI). A low process noise variance is employed by default to ensure intensive smoothing, while the process noise variance is momentarily increased if the correction vector is found to approach the limit of permitted frame shifts.

II. KALMAN FILTERING

The Kalman filter is implemented to provide an estimate to the state of a discrete-time process that is defined in the recursive form of a linear dynamic system [18].

$$x(t+1) = F * x(t) + w(t) \quad (1)$$

The state of the system is linearly related to the previous state through the matrix F , and $w(t)$ is used to denote the process noise. The Kalman filter operates using observations of all or some of the state variables, defined by the observation system:

$$y(t) = H * x(t) + v(t) \quad (2)$$

The observation variables are linearly related to the state variables through the matrix H , while $v(t)$ is used to denote the measurement noise. By definition of the Kalman filter, process and measurement noise are assumed to be independent of each other, white, having normal probability distributions $w \sim N(0, Q)$ and $v \sim N(0, R)$.

The Kalman filter predicts the process state using a form of feedback control: first the process state is obtained from previous state variables through the linear dynamic system defined by (1), then a feedback is attained from the measurement input. Thus the Kalman filtering process

is divided into two parts: the prediction stage (time update equations) and the correction stage (measurement update equations). The time update equations project the current state and error covariance estimates forward (in time) to obtain the a priori state estimates. The measurement update equations are responsible for the feedback, by incorporating a new measurement into the a priori estimates to obtain the improved (tuned) a posteriori estimates.

The prediction is accomplished through the linear dynamic system by $\hat{x}_{n+1}^- = F * \hat{x}_n^-$, where \hat{x}_{n+1}^- is the a priori state estimate for the next state (image frame $n+1$) and \hat{x}_n^- is the a posteriori state estimate of the current state (frame n). At the prediction stage, the a priori estimate error covariance, that is defined as $P_n = E[(x_n - \hat{x}_n^-)(x_n - \hat{x}_n^-)^T]$, for frame $n+1$ is obtained recursively from $P_{n+1}^- = F P_n F^T + Q$, where P_n is the a posteriori estimate error covariance for frame n .

The correction is accomplished by computing the gain matrix $K_n = P_n^- H^T (H P_n^- H^T + R)^{-1}$ that is used to compute the a posteriori state from the a priori estimates by $\hat{x}_n = \hat{x}_n^- + K_n (y_n - H \hat{x}_n^-)$. The a posteriori estimate error covariance is computed similarly from $P_n = (I - K_n H) P_n^-$.

After each time and measurement update pair, the process is repeated with the previous a posteriori estimates used to predict the new a priori estimates. Thus, the Kalman filter recursively conditions the current estimate on all of the past measurements. This recursive nature of the Kalman filter enables practical real-time implementation.

To employ the Kalman filter, the state and observation equations are constructed as in (1) and (2), to define the dynamic process and relate the corresponding measured inputs. The respective noise variances have to be set according to process and measurement noise characteristics to enable optimal operation. Then the model can simply be plugging into a generic form of the Kalman filter [18], that carries out the resulting algebra (prediction and correction stages), to obtain the state estimates for each instance.

Kalman filtering has been employed in various fields of image processing, such as video restoration [19]. In cases where the process to be estimated, or the measurement relationship is non-linear, a Kalman filter that linearizes about the current mean and covariance, referred to as the extended Kalman filter can be employed. For example, the use of an extended Kalman filter for real-time

estimation of long-term 3-D motion parameters for model based coding has been presented in [20], and the use of an extended Kalman filter for object tracking has been presented in [21].

For the Kalman filter based image stabilization system, a linear global camera motion model is used, and the observations are also linearly related to the process state. Therefore, the standard Kalman filter can be employed satisfactorily for the smoothing of global motion, with no need for the more complex extended Kalman filter.

III. CONSTANT VELOCITY CAMERA MODEL

In order to ensure a smooth frame transition, the global camera motion is modelled as a constant velocity motion process. According to defined process noise characteristics, the Kalman filter will actually allow variations in the camera velocity to agree with measured global frame displacements, instead of enforcing constant velocity frame displacements. The important feature is that the velocity variations are driven to be smooth.

Any real-time motion estimation technique, for instance gray-coded bit-plane matching as demonstrated in [6], can be employed to obtain interframe global motion vectors. The absolute position of an image frame is defined as the absolute displacement with respect to the first frame of that particular scene shot. The absolute frame position of any frame can be obtained from the accumulation of all former interframe differential global motion vectors. It is also possible to track absolute frame position recursively, by adding the interframe motion vector of a particular frame to the absolute position of the previous frame.

In order to construct the state and measurement equations for the camera motion system, the absolute frame position can be assigned to correspond to the instantaneous absolute camera position. For the state equations of the constant velocity camera motion model (CVCMM), the position of a frame is defined as the position of the previous frame plus the constant camera velocity, in each direction. Denoting x_x and x_y as the absolute horizontal and vertical positions in the image plane respectively, and v_x and v_y as the horizontal and vertical camera velocities, the CVCMM model state system is expressed as

$$\begin{bmatrix} x_x(n) \\ x_y(n) \\ v_x(n) \\ v_y(n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_x(n-1) \\ x_y(n-1) \\ v_x(n-1) \\ v_y(n-1) \end{bmatrix} + [w] \quad (3)$$

The measurement system is constructed to take the absolute frame positions as input to the system, and thus the observation equations are expressed in the form of

$$\begin{bmatrix} mx_x(n) \\ mx_y(n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_x(n) \\ x_y(n) \\ v_x(n) \\ v_y(n) \end{bmatrix} + [v] \quad (4)$$

Having constructed the state-transition and observation equations, it is important to set process and measurement noise variance values reasonably, to ensure appropriate execution as operation of the Kalman filter directly depends on the values of the process noise variance Q and the measurement variance R .

The measurement noise variance has shown to determine how fast the Kalman filter reacts to the observations. A relatively large measurement noise variance results in loosing track of intentional camera movements as the Kalman filter is 'slow' to believe the measurements, while a relatively small value may cause inappropriate stabilization as the Kalman filter will follow fluctuations because it is 'very quick' to believe the measurements. It is possible to set the measurement noise variance up front, according to presumed jitter characteristics, a measurement noise variance value reasonably close to the actual jitter variance is found to enable adequate operation throughout the stabilization process.

The process noise variance on the other hand, has shown to condition the adaptability of the Kalman filter to changes in the direction or speed of intentional camera motion. A relatively small process noise variance restrains the Kalman filter to constant velocity operation, in which case intentional movements can not be tracked accurately if camera motion dynamics change. A relatively large process noise variance causes the Kalman filter to follow measured global motion values closely, in which case it is possible that the filtered output follows low-frequency jitter that actually needs to be removed. The utilisation of an intermediate process noise variance results in compromises in stabilization performance as well as deliberate camera movement preservation. Furthermore intensive stabilization may cause the correction vector to exceed the limit of permitted frame shifts and cause a blank region to be displayed within the visible image frame. Instead of assigning the process noise variance to a predetermined value, it is proposed in this paper to set it adaptively according to previous correction vectors. A small process noise variance is employed as default to enable profound stabilization in general, however the process noise variance is increased if the correction vector approaches the limit of permitted frame shifts.

The state-transition and observation matrices, and the process and measurement noise variances are plugged into the generic Kalman filter equations to obtain a real-time estimate for the horizontal and vertical camera positions

that make up the state variables. As the constant velocity camera motion model directly smoothens the absolute frame displacements, the correction vector for any frame is obtained from the difference of the Kalman filtered and the original absolute frame position:

$$V_{cor}(n) = X_{Klm}(n) - X_{act}(n) \quad (5)$$

IV. RESULTS AND DISCUSSION

Figure 1 shows a sample frame of the off-road sequence (source: <http://www.cfar.umd.edu/~sirohey/ZD.html>). The corresponding absolute global frame displacements are displaced in Figure 2. The sequence originates from a camera mounted on a ground vehicle navigating on an unsteady surface, advancing parallel to the vehicle in the scene. The regularly increasing horizontal displacement projects the motion of the camera vehicle, while the irregular vertical displacements reflect the surface structure.



Figure 1. Sample frame of the off-road image sequence

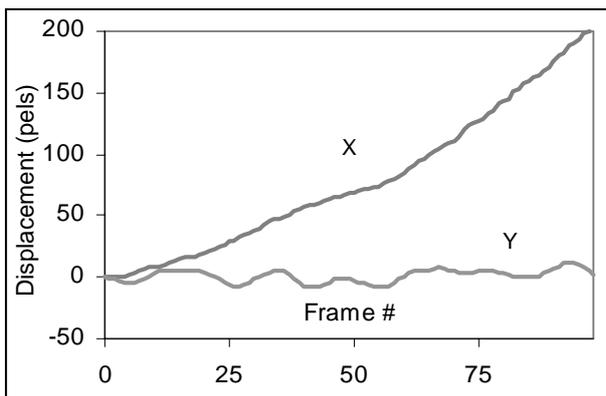


Figure 2. Absolute raw frame displacements of the off-road image sequence

It is seen from Figure 2 that the standard deviation (σ) of absolute frame positions from the long-term movement trajectory is particularly low for the horizontal direction,

and slightly larger for the vertical one. The standard deviation for the horizontal direction is found to be about 1.5 pixels, while the standard deviation for the vertical is found to be about 5 pixels. Ideally the measurement noise variance $R (= \sigma^2)$ would be set to a value in the range of $R \sim 3$ for the horizontal and $R \sim 25$ for the vertical. However, as standard deviation of jitter might be larger for other sequences, the utilisation of higher measurement noise variances in the range of $R \sim 100$ (corresponding to a jitter standard deviation of 10 pixels) can be employed successfully to provide robustness to varying jitter characteristics.

The Kalman filtered frame positions for the off-road sequence, having employed a relatively large process noise variance of $Q=0.1$ is given in Figure 3. It is seen that the positions of the stabilized sequence closely follow the original ones. Although this has the advantage of small correction vectors, that do not exceed the permitted frame shift limit, the drawback is that the short-term fluctuations encountered in the vertical are not stabilized.

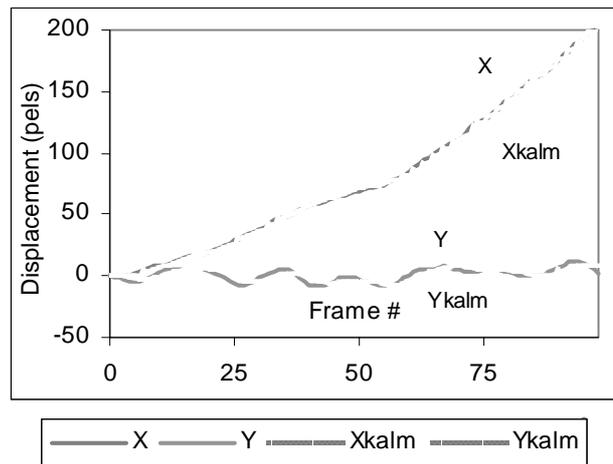


Figure 3. Raw and Kalman filtered ($Q=0.1, R=100$) frame displacements of the off-road image sequence

The Kalman filtered frame positions for the off-road sequence, having employed a relatively low process noise variance of $Q=0.0001$ is given in Figure 4. Stabilization intensity is increased, as the Kalman filter forces the system to constant velocity operation more tightly, however large correction vectors are encountered in the horizontal direction, inclined to exceed the limit.

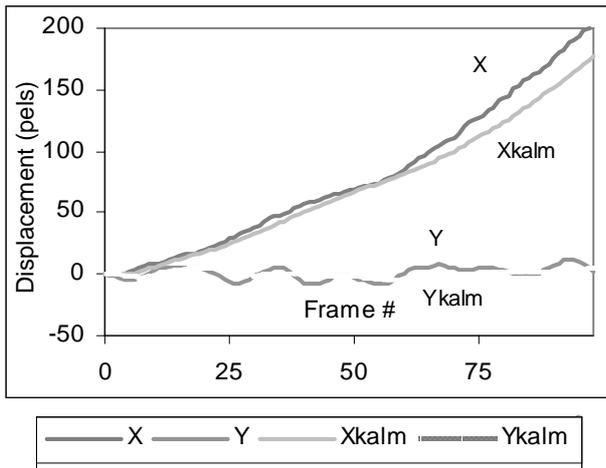


Figure 4. Raw and Kalman filtered ($Q=0.0001$, $R=100$) frame displacements of the off-road image sequence.

It is seen that a relatively low process noise variance is preferable in the stabilization point of view as even short-term fluctuations are smoothened. However if the motion dynamics shows a change from the constant velocity motion, the stabilizer is late to react to these changes at low process noise variance as the Kalman filter enforces stability of the constant velocity camera motion process. For instance, for the off-road sequence the horizontal camera velocity is about constant until frame sixty, at which point the camera velocity changes due to vehicle motion. At a low process noise variance the stabilized sequence follows the original sequence closely to this point, but is late to react to the changes in motion dynamics causing the stabilized sequence to stay at the former velocity, resulting in large correction vectors. Therefore, correction vectors of up to 30 pixels are encountered in the horizontal direction for the low process noise variance case.

Figure 5 shows an example construction, for the process noise variance of the Kalman filtering process being adaptively changed according to the magnitude of the previous correction vector. A low process noise variance is used if the Kalman filtered positions are close to the original ones, and larger process noise variances are employed if the filtered positions deviate from the original case. In this case, the limit of permitted frame shifts is set to 16 pixels, and the process noise variance is set according to the last correction vector computed from equation (5) in the form of:

$$\begin{aligned}
 & \text{if } V_c(n-1) > 14 & Q_n &= 0.5 \\
 & \text{else if } V_c(n-1) > 11 & Q_n &= 0.2 \\
 & \text{else if } V_c(n-1) > 8 & Q_n &= 0.1 \\
 & \text{else} & Q_n &= 0.0001
 \end{aligned} \tag{6}$$

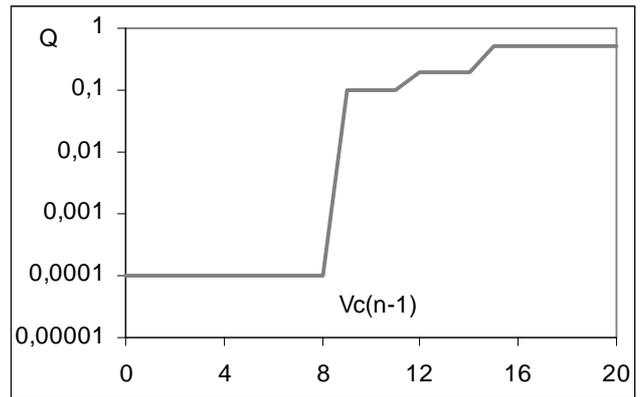


Figure 5. Process noise variance values set adaptively according to the last correction vector.

The resultant absolute frame positions for the Kalman filtered off-road sequence with process noise variance set adaptively according to (6) are displayed in Figure 6. It is clearly seen that in the horizontal direction, the Kalman filter adjusts to original position signal if the difference (the correction vector magnitude) increases, realised by the increased process noise variance. The maximum correction vector magnitude has been found to be 13 pixels for the adaptive process noise stabilization process, which is well within the 16 pixel constraint set to be the limit of permitted frame shifts.

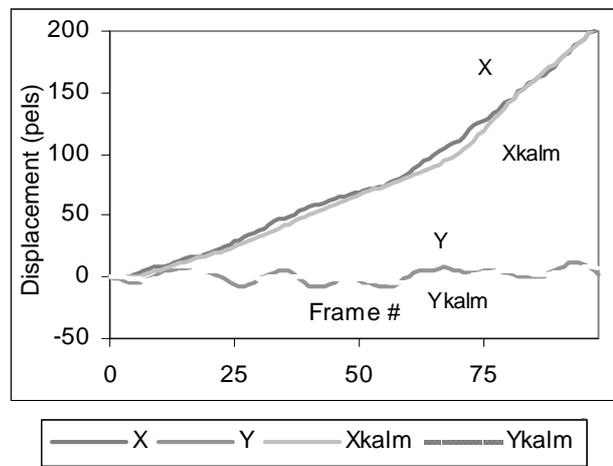


Figure 6. Raw and Kalman filtered (Q : adaptively set according to (6), $R=100$) frame displacements of the off-road image sequence.

V. CONCLUSION

Image sequence stabilization by Kalman filtering using a constant velocity camera motion model with adaptive process noise variance has been presented. It is shown that by changing the process noise according to the magnitude of previous correction vectors, it is possible to combine the superior stabilization performance gained at low process noise variance with the improved adjustment capability of high process noise variance. The typically low process noise variance ensures increased stabilisation

intensity enabling the cancellation of short-term fluctuations. With adaptation the process noise variance is raised in cases where the correction vector amplitude increases towards the limit of permitted frame shifts, to ensure that the correction vector stays within the limit to avoid blank regions to be displayed in the visible frame. The simple recursive implementation of the Kalman filter and simple control of the process noise variance enable real-time utilisation of the proposed image sequence stabilization system.

REFERENCES

1. M. Oshima, Y. Hayashi, S. Fujiko, T. Inaji, H. Mitani, J. Kajino, K. Ikeda, and K., Komoda: "VHS camcorder with electronic image stabilizer", *IEEE Trans. Consum. Electron.*, 35, (4), pp. 749-757, 1989.
2. T. Vieville, E. Clergue, and P.E.D.S. Facao: "Computation of ego-motion and structure from visual and internal sensors using the vertical cue", *Proceedings of IEEE Int. Conf. On Computer Vision, Berlin*, pp. 591-598, 1993.
3. K. Sato, S. Ishizuka, A. Nikami, and M. Sato: "Control techniques for optical image stabilizing system", *IEEE Trans. Consum. Electron.*, 39, (3), pp. 461-465, 1993.
4. K. Uomori, A. Morimura, and H. Ishii: "Electronic image stabilization systems for video cameras and VCRs", *J. Soc. Motion Pict. Telev. Eng.*, 101, (2), pp. 66-75, 1992.
5. S.J. Ko, S.H. Lee, and K.H. Lee: "Digital image stabilizing algorithms based on bit-plane matching", *IEEE Trans. Consum. Electron.*, 44, (3), pp. 617-622, 1998.
6. S.J. Ko, S.H. Lee, S.W. Jeon, and E.S. Kang: "Fast digital image stabilizer based on gray-coded bit-plane matching", *IEEE Trans. Consum. Electron.*, 45, (3), pp. 598-603, 1999.
7. J.K. Pail, Y.C. Park, and D.W. Kim: "An adaptive motion decision system for digital image stabilizer based on edge pattern matching", *IEEE Trans. Consum. Electron.*, 38, (3), pp. 607-615, 1992.
8. P. Burt, and P. Anandan: "Image stabilization by registration to a reference mosaic", *Proc. of ARPA Image Understanding Workshop, Monterey, CA*, pp. 425-434, 1994.
9. M. Hansen, P. Anandan, K. Dana, G. Van Der Wal, P.J. Burt: "Real-time scene stabilization and mosaic construction", *Proc. of ARPA Image Understanding Workshop, Monterey, CA*, pp. 457-465, 1994.
10. C. Morimoto, and R. Chellappa: "Fast electronic digital image stabilization for off-road navigation", *Real-Time Imaging*, 2, pp. 285-296, 1996.
11. S. Ertürk, and T.J. Dennis: "Image sequence stabilisation based on DFT filtering", *IEE Proc. On Vision Image and Signal Processing*, 147, (2), pp. 95-102, 2000.
12. Y.S. Yao, P. Burlina, and R. Chellappa: "Electronic image stabilization using multiple visual cues", *Proc. of IEEE Int. Conf. on Image Processing, Washington DC*, 1995
13. Z. Duric and A. Rosenfeld: "Image sequence stabilization in real time", *Real-Time Imaging*, 2, pp. 271-284, 1996.
14. M. Irani, B. Rousso, and S. Peleg: "Recovery of ego-motion using image stabilization", *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition, Seattle*, pp. 454-460, 1994.
15. Y.S. Yao, and R. Chellappa: "Selective stabilization of images acquired by unmanned ground vehicles", *IEEE Trans. Robotics and Automation*, 13, (5), pp. 693-708, 1997.
16. A. Engelsberg, and G. Schmidt: "A comparative review of digital image stabilising algorithms for mobile video communications", *IEEE Trans. Consum. Electron.*, 45, (3), pp. 591-597, 1999.
17. S. Ertürk: "Image sequence stabilisation based on Kalman filtering of frame positions", *Electronics Letters*, 37, (20), pp. 1217-1219, 2001.
18. Brown, R.G. & Hwang, P.Y.C. (1992). *Introduction to random signals and applied Kalman filtering* (2.nd ed.), John Wiley & Sons.
19. Patti, A.J., Tekalp, A.M. & Sezan, M.I. (1998). A new motion-compensated reduced-order model Kalman filter for space-varying restoration of progressive and interlaced video. *IEEE Trans. Image Processing*, 7: 543-554.
20. Smolic, A., Makai, B. & Sikora, T. (1999). Real-time estimation of long-term 3-D motion parameters for SNHC Face Animation and Model-Based coding applications. *IEEE Trans. Circuits and Syst. for Video Technology*, 9: 255-263.
21. Foresti, G.L. (1999). Object Recognition and tracking for remote video surveillance. *IEEE Trans. Circuits and Syst. for Video Technology*, 9: 1045-1062.