# CONSTRAINT TRANSLATIONAL AND ROTATIONAL MOTION FILTERING FOR VIDEO STABILIZATION

*Marius Tico, Markku Vehvilainen*

Nokia Research Center
P.O.Box 100, FI-33721 Tampere, Finland
email: tico@ieee.org

## ABSTRACT

In this paper we propose a novel motion filtering approach that takes into consideration the existence of certain system constraints with respect to the amount of the corrective rotational and translational motions that can be applied on each video frame for stabilization. The interdependence between rotational and translational constraints is considered, and a modified Kalman filtering algorithm is used in order to obtain a smooth stabilized motion under the given constraints. The experimental results reveal that the proposed filtering approach improves the stabilization performance in the presence of the system constraints.

## 1. INTRODUCTION

The ongoing development and miniaturization of consumer devices that have video acquisition capabilities increases the need for robust and efficient video stabilization solutions. Video stabilization objective is to remove the effect of unwanted motion fluctuations from video data. This can be achieved by applying a certain amount of corrective motion displacement onto each video frame, such that to cancel the effect of high frequency fluctuations (jitter) caused by unwanted camera motions.

A video stabilization system comprises three components: global motion estimation, motion filtering, and motion compensation. The global motion comprises two components: the user intended motion, and the unwanted motion. The objective of motion filtering operation is to distinguish between these two motion components such that to allow subsequent compensation only for the undesired motion. For this, it is typically assumed that the intended motion component is smooth, such that it can be calculated by low-pass filtering the estimated global motion.

In [1] the authors proposed to low-pass filter the camera motion trajectory in Fourier domain. The solution provides a smooth stabilized motion and it can be applied for off-line stabilization of a recorded video sequence. Unfortunately, the solution is unsuitable for a real-time implementation on a typical consumer device due to its large memory requirements needed to store several frames of the input video sequence. For real-time implementation a causal low-pass filter is preferred in order to reduce the memory requirements to a minimum.

First order IIR (infinite impulse response) low pass filtering system, known as Motion Vector Integration (MVI), is used in [2], and analyzed also in [3]. The main drawback of MVI consists of its tradeoff between smoothness of the resulted stabilized motion and the delay in reaction with respect to changes in the intended motion. The damping coefficient of the filter must be selected such that to cope with this tradeoff. Second order IIR filter, inertial filter, has been proposed in [4], as an attempt to reduce the phase delay, and hence to enhance the ability to follow any intended changes in the camera motion. Kalman filtering have been used for video stabilization in [5, 6, 7], and it has been proven to be a simple and robust solution for on-line video stabilization implementations.

The motion compensation process of the video stabilization system consists of geometrically transforming each video frame such that to cancel the effect of unwanted motion. As a result of such geometrical transformation, some parts of the stabilized video frame could become undefined, being placed outside the captured image area. In order to prevent this phenomenon one could use a smaller frame size at the output of the stabilization system than at the input. In this way, the stabilized frame is obtained by cropping a certain region from the larger input frame delivered by the image sensor. The difference in size between the input and the output image frames of the stabilization system determines certain constraints with resect to the rotational and translational corrective motions that could be applied by the stabilization system. Any larger corrective motions would result into an incomplete output frame, as long as part of it falls outside the boundaries of the sensed image frame.

In this paper we address the problem of motion filtering taking also into consideration the limitation imposed by the stabilization systems with respect to the amount of corrective displacement that can be applied on each video frame. Thus, we propose to incorporate the system constraint into the Kalman filtering procedure, as an additional state update equation. In this way, the stabilized motion trajectory is determined as the solution of a constraint optimization problem.

The remaining part of the paper is organized as follows. A constraint Kalman filtering algorithm is introduced in Section 2. The proposed motion filtering algorithm, under the system constraints on corrective motions, is introduced in Section 3, and it is validated by a series of experiments presented in Section 4. Finally, some conclusions are presented in Section 5 of the paper.

## 2. CONSTRAINT KALMAN FILTERING ALGORITHM

Let $z_n$ denotes one of the motion parameters (i.e. translation or rotation) estimated based on the $n$-th frame of the video sequence. We can write that

$$z_n = s_n + u_n, \tag{1}$$

where $s_n$ and $u_n$ stand respectively for the intended and unwanted components of the motion parameter. At every moment $n$, a correction of $-u_n$ is needed in order to stabilize the current frame. In practice, the value of this corrective displacement is constraint by the system to a given interval $[d_{min}, d_{max}]$. Our objective can be thereby formulated as determining a smooth trajectory for the intended motion component ($s_n$), under the constraint that at any moment $n$,

$$d_{min} \leq -u_n \leq d_{max}. \tag{2}$$

A state space representation model for the motion parameter can be assumed as follows

$$\begin{aligned} \mathbf{x}_n &= \mathbf{A}\mathbf{x}_{n-1} + \mathbf{b}e_n, \\ z_n &= \mathbf{c}^T\mathbf{x}_n + u_n, \end{aligned} \tag{3}$$

where $e_n$ and $u_n$ are the process and measurement noise terms that are assumed zero mean Gaussian distributed with variances $\sigma_e^2$ and $\sigma_u^2$ respectively. The process matrix $\mathbf{A}$ has size $K \times K$, and the vectors $\mathbf{c}$ and $\mathbf{b}$ are of size $K \times 1$. The state $\mathbf{x}_n$ is a $K \times 1$ vector from which the intended motion $s_n$ can be extracted by $s_n = \mathbf{c}^T\mathbf{x}_n$.

Kalman filtering procedure provides the optimal estimate $\hat{\mathbf{x}}_n$, (and ultimately the optimal estimate of $s_n$), based on the assumed model (3). However, in this form, the model is incomplete since it does not take into consideration the constraint (2), which, in terms of the state space representation (3) is given by

$$\|z_n - \mathbf{c}^T\mathbf{x}_n - D\|^2 \leq d^2, \tag{4}$$

where $D = (d_{min} + d_{max})/2$, and $d = (d_{max} - d_{min})/2$. Whenever this constraint applies the solution provided by Kalman procedure cannot be used directly. Our objective here is to derive a modified Kalman filtering procedure that incorporates the constraint in an optimal manner.

Let $\hat{\mathbf{x}}_n$ denotes the Kalman filter estimate of the true state $\mathbf{x}_n$, in the absence of any constraint. In accordance to the Kalman filter theory (see also [8]), the posterior PDF (probability density function) of the true state given all observations is given by

$$p(\mathbf{x}_n|z_{n:1}) \sim \mathcal{N}(\hat{\mathbf{x}}_n, \mathbf{P}_n), \tag{5}$$

where $\mathcal{N}(\mu, \Sigma)$ stands for multivariate Gaussian PDF of mean $\mu$ and covariance matrix $\Sigma$, and $\mathbf{P}_n$ stands for the estimate of the true state covariance matrix at step $n$

$$\mathbf{P}_n = E\left[(\mathbf{x}_n - \hat{\mathbf{x}}_n)(\mathbf{x}_n - \hat{\mathbf{x}}_n)^T\right]. \tag{6}$$

The estimate $\hat{\mathbf{x}}_n$ maximizes thereby the posterior PDF (5), or, similarly, it minimizes the following objective function

$$J'(\mathbf{x}_n) = (\mathbf{x}_n - \hat{\mathbf{x}}_n)^T\mathbf{P}_n^{-1}(\mathbf{x}_n - \hat{\mathbf{x}}_n). \tag{7}$$

The estimate is optimal as long as it satisfies (4). Otherwise, the optimal solution $\tilde{\mathbf{x}}_n$, can be found by employing Lagrange multipliers and treating the constraint (4) as an equality constraint. Renouncing to the subscript $n$, in order to simplify the notations, we can write an objective function to be minimized by the constraint estimate $\tilde{\mathbf{x}}$, as follows

$$\begin{aligned} J(\mathbf{x}) &= (\mathbf{x} - \hat{\mathbf{x}})^T\mathbf{P}^{-1}(\mathbf{x} - \hat{\mathbf{x}}) \\ &+ \lambda\left[(\mathbf{c}^T\mathbf{x} - q)^T(\mathbf{c}^T\mathbf{x} - q) - d^2\right], \end{aligned} \tag{8}$$

where $\lambda$ is a positive Lagrange multiplier, and $q = z - D$. Differentiating (8) with respect to $\mathbf{x}$ and equating the result with zero we have

$$\tilde{\mathbf{x}} = (\mathbf{P}^{-1} + \lambda\mathbf{c}\mathbf{c}^T)^{-1}(\lambda\mathbf{c}q + \mathbf{P}^{-1}\hat{\mathbf{x}}), \tag{9}$$

which, replaced into the constraint equation yields the following value for the Lagrange multiplier

$$\lambda = (\mathbf{c}^T\mathbf{P}\mathbf{c})^{-1}\frac{|\mathbf{c}^T\hat{\mathbf{x}} - q| - d}{d}. \tag{10}$$

Finally, including (10) in (9), and performing some calculus we obtain the expression of the update equation for the state estimate

$$\tilde{\mathbf{x}} = \hat{\mathbf{x}} + \mathrm{sign}(q - \mathbf{c}^T\hat{\mathbf{x}})(|q - \mathbf{c}^T\hat{\mathbf{x}}| - d)\mathbf{P}\mathbf{c}(\mathbf{c}^T\mathbf{P}\mathbf{c})^{-1}. \tag{11}$$

The algorithm presented in Fig. 1, extends the Kalman filtering procedure by incorporating the state update equation (11). In addition, the first line of the algorithm is updating also the measurement noise variance ($\sigma_u^2$) such that to reduce the probability of exceeding the system constraint.

$CKF(z_n, d_n, D_n)$
 $\sigma_u = d_n/3$
 $\mathbf{x}_{n|n-1} = \mathbf{A}\mathbf{x}_{n-1}$
 $\hat{\mathbf{P}} = \mathbf{A}\mathbf{P}\mathbf{A}^T + \sigma_e^2\mathbf{b}\mathbf{b}^T$
 $\mathbf{g} = \hat{\mathbf{P}}\mathbf{c}\left(\mathbf{c}^T\hat{\mathbf{P}}\mathbf{c} + \sigma_u^2\right)^{-1}$
 $\mathbf{P} = \left(\mathbf{I}_K - \mathbf{g}\mathbf{c}^T\right)\hat{\mathbf{P}}$
 $\mathbf{x}_n = \mathbf{x}_{n|n-1} + \mathbf{g}\left(z_n - \mathbf{c}^T\mathbf{x}_{n|n-1}\right)$
 $u_n = z_n - \mathbf{c}^T\mathbf{x}_n$
 $w_n = u_n - D$
 if $|w_n| > d_n$ then
  $\mathbf{x}_n = \mathbf{x}_n + \mathtt{sign}(w_n)(|w_n| - d_n)\mathbf{P}\mathbf{c}(\mathbf{c}^T\mathbf{P}\mathbf{c})^{-1}$
  $u_n = z_n - \mathbf{c}^T\mathbf{x}_n$
 return $-u_n$

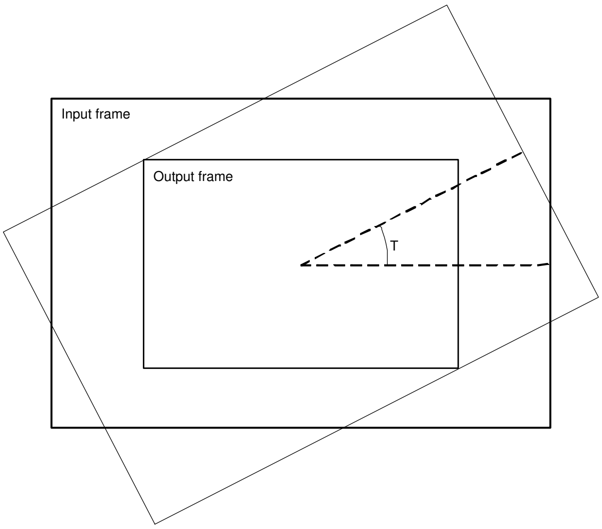
Figure 1: The proposed constraint Kalman filtering procedure.

In contrast with this solution, a trivial approach would consists of running unconstraint Kalman filtering procedure, following to disregard the resulted correction whenever it exceeds the system constraint. In such cases, the correction recommended by the filtering procedure is simply truncated to the maximum corrective value allowed by the system.

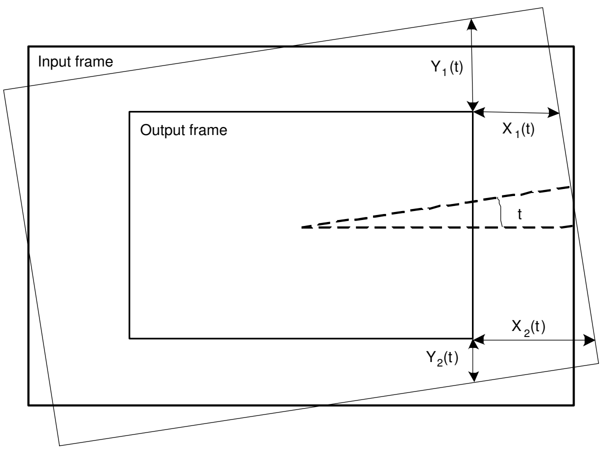
## 3. THE ROTATION AND TRANSLATION MOTION FILTERING

In our work we consider a motion model composed of translation and rotation. The three motion parameters estimated from the input video data are as follows: $x_n^e$ translation along the horizontal axis, $y_n^e$ translation along the vertical axis, and $t_n^e$ rotation around the optical axis of a camera. The corresponding corrective displacements for each parameter are denoted respectively by $x_n^c$, $y_n^c$, and $t_n^c$.

The system constraints on the three motion parameters are illustrated in Fig.2. The figure shows the input and output frames of the stabilization algorithm. The rotational constraint (i.e. $t \in [-T, T]$) is the first one applied, and hence its value is independent of translational corrections. On the other hand, the translational constraints are dependent of the rotational correction decided in the first step of the algorithm, as illustrated in Fig. 2 (b). These constraints are calculated such that to maintain the output image frame inside the rotated input frame. According to the example shown in the figure, the horizontal and vertical constraints are respectively $X(t) = min\{X_1(t), X_2(t)\}$, and $Y(t) = min\{Y_1(t), Y_2(t)\}$.




(a)




(b)

Figure 2: The constraints on rotational correction (a), and translational correction (b).

The motion filtering algorithm used in our work is presented in Fig. 3. The algorithm runs three different Kalman filters in order to smooth the rotational and translational motion parameters.

## 4. EXPERIMENTAL RESULTS

In this section we present a series of experiments meant to demonstrate the effectiveness of the proposed modification to normal Kalman filtering procedure. For this purpose we implemented the motion filtering procedure proposed in [5]. This procedure assumes a model of camera motion characterized by a constant velocity between moments of intentional motion changes. Each state space model is using a

---

$t_n^c = \texttt{CKF1}(t_n^e, T, 0)$
`Calculate constraints` $X(t_n^c)$ `and` $Y(t_n^c)$
$x_n^c = \texttt{CKF2}(x_n^e, X(t_n^c), 0)$
$y_n^c = \texttt{CKF3}(y_n^e, Y(t_n^c), 0)$

---

Figure 3: The proposed constraint filtering procedure for rotational and translational jitter.

$2 \times 1$ state vector whose components represent respectively the smoothed motion parameter value $s_n$, and the velocity along the corresponding direction. The remaining parameters of the state space model are:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \ \mathbf{c} = [1\ 0]^T, \text{ and } \mathbf{b} = [1\ 1]^T. \qquad (12)$$

The following experiments have been carried out on a video sequence captured from a moving car. As shown in Fig.4, the sequence exhibits a significant intended motion along horizontal, combined with a less significant intended displacement along the vertical position. Unwanted motion is present along both horizontal and vertical directions, having a more noticeable effect along the vertical direction.

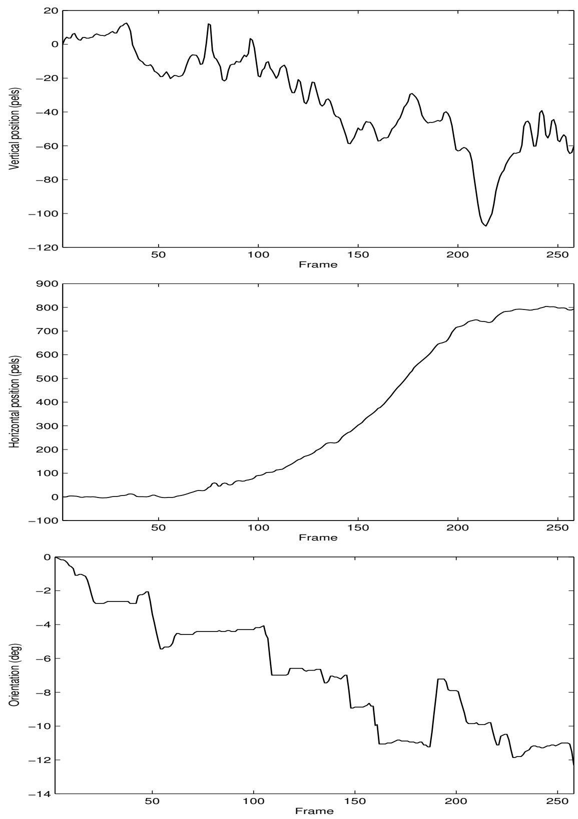



Figure 4: The frame positions and orientations for the experimental video sequence.

The degree of freedom needed to compensate for unwanted motion is ensured by using a smaller frame size at the output of the stabilization procedure than at the input. In all our experiments we selected the output frame size with $2d$

pixels smaller than the input frame size along the horizontal and vertical directions.

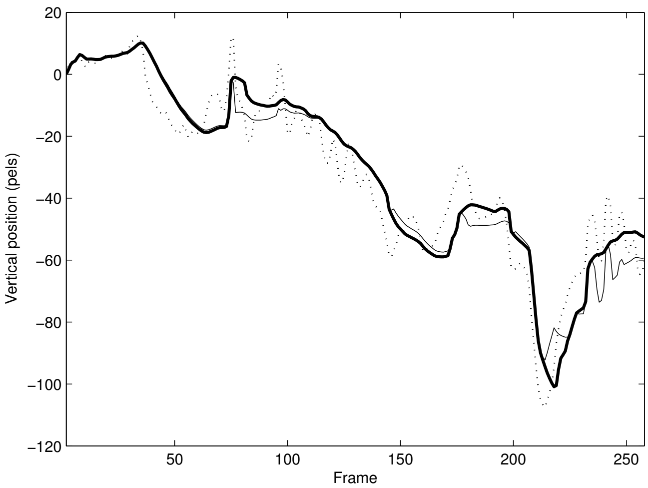



Figure 5: The filtered frame positions for constraint $d = 16$ pels. The figure shows the original motion trajectory (dotted line), the trivial solution (thin line), and the solution deliver by the proposed algorithm (thick line).

Fig. 5 shows an example of motion filtered trajectories obtained under a constraint of $d = 16$ pixels. The figure reveals that the resulted motion trajectory is much smoother when using the proposed approach than in the case a trivial solution for the incorporation of the system constraint is adopted.

The jitter attenuation (in decibels) has been used to evaluate the stabilization performance. This parameter is calculated as the ratio between the unwanted motion energies present in a motion trajectory before and after filtering. The energy of the unwanted motion fluctuations (jitter) present in one such trajectory is estimated by the variance of the high-pass filtered version of the corresponding frame position signal, where the high pass filtering is carried out in the frequency (DFT) domain.

Table 1 shows the jitter attenuation achieved for different values of the constraint $d$, when using either the proposed solution (a), or a trivial saturation of the motion correction to the maximum value allowed in the system (b). The test sequence exhibits a small rotational jitter which is rarely challenging the rotational constraint imposed by the system. Consequently in all these simulations the rotational jitter attenuation is the same for both cases. However, the rotational correction imposes different constraints for the horizontal and vertical corrections. All simulations have been performed using the same parameters for the state space model, such that the differences in performance are caused only by the strategy used to incorporate the system constraint into the filtering procedure. The results shown in this table reveal that employing the proposed solution in the presence of a system constraint improves the stabilization performance.

| d | 8 | 16 | 24 | 32 | 64 |
|---|---|---|---|---|---|
| Horizontal | | | | | |
| (a) | 4.2 | 5.2 | 5.2 | 6.0 | 8.2 |
| (b) | 1.5 | 0.5 | 1.1 | 2.6 | 2.9 |
| Vertical | | | | | |
| (a) | 4.6 | 8.1 | 11.0 | 15.1 | 21.1 |
| (b) | 2.4 | 4.8 | 8.8 | 15.0 | 21.1 |
| Rotational | | | | | |
| (a) | 12.0 | 15.2 | 16.4 | 17.1 | 25.0 |
| (b) | 12.0 | 15.2 | 16.4 | 17.1 | 25.0 |

Table 1: Jitter attenuation (in dB) under different system constraints ($d$) imposed on maximum corrective displacement: (a) using the proposed filtering procedure, and (b) applying a trivial approach which saturates the amount of corrective displacement to the nearest value accepted by the system.

## 5. CONCLUSIONS

In this paper we introduced a new method of motion filtering for video stabilization. Both translational and rotational jitter corrections are considered, and they are corrected under certain system constraints. The proposed approach is based on a modified Kalman filtering algorithm that includes an additional state update equation which ensures that the updated state maximizes the posterior PDF, under the imposed constraint. The ability of the proposed filtering procedure to cope with the constraint on motion correction magnitude, is demonstrated through a series of video stabilization experiments.

## REFERENCES


[1] S. Erturk and T.J. Dennis, "Image sequence stabilization based on DFT filtering," *IEE Proc. On Vision Image and Signal Processing*, vol. 147, no. 2, pp. 95–102, 2000.

[2] S. J. Ko, S. H. Lee, and K. H. Lee, "Digital image stabilizing algorithms based on bit-plane matching," *IEEE Transaction on Consumer Electronics*, vol. 44, no. 3, pp. 617–622, 1998.

[3] Sarp Erturk, "Image sequence stabilization: Motion vector integration (MVI) versus frame position smoothing (FPS)," in *Proc. of the 2nd International Symposium on Image and Signal Processing and Analysis*, 2001, pp. 266–271.

[4] Jesse S. Jin, Zhigang Zhu, and Guangyou Xu, "A Stable Vision System for Moving Vehicles," *IEEE Transaction on Intelligent Transportation Systems*, vol. 1, no. 1, pp. 32–39, 2000.

[5] S. Erturk, "Image sequence stabilization based on Kalman filtering of frame positions," *Electronics Letters*, vol. 37, no. 20, pp. 1217–1219, 2001.

[6] Andrew Litvin, Janusz Konrad, and William C. Karl, "Probabilistic video stabilization using Kalman filtering and mosaicking," in *Proc. of SPIE Electronic Imaging*, 2003, vol. 5022, pp. 663–674.

[7] S. Erturk, "Digital image stabilization with sub-image phase correlation based global motion estimation," *IEEE Transaction on Consumer Electronics*, vol. 49, no. 4, pp. 1320–1325, 2003.

[8] Dan Simon and Tien Li Chia, "Kalman filtering with state equality constraints," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 1, pp. 128–136, 2002.