The 19-th IEEE Conference in Israel, Jerusalem, November 1996.

RECURSIVE OPTICAL FLOW ESTIMATION - ADAPTIVE FILTERING APPROACH

M. Elad and A. Feuer
The Technion - Israel Institute of Technology
The Electrical Engineering Department - Haifa 32000, Israel

Abstract

This paper presents a new approach towards the problem of recursive optical flow estimation from image sequences based on the differential framework proposed by Horn & Schunck. We show that gain is achieved both from computational and accuracy points of view when treating the estimation task recursively. Incorporation of the temporal axis into the estimation process is done by combining a temporal smoothness assumption on the optical flow. RLS and LMS recursive optical flow estimation algorithms are derived and tested.

1. Introduction

Optical flow is the displacement field related to each of the pixels in an image sequence. Such displacement field results from the apparent motion of the image brightness in time. Estimating the optical flow is a fundamental problem in low level vision, and can undoubtedly serve many applications in image sequence processing. There are many different methods to estimate the optical flow. This paper focuses on a generalization of the method proposed by Horn & Schunck, which is a differential based method [1].

Most of the algorithms for the estimation of optical flow concentrate on the goal of estimating the motion field between succeeding images in a sequence, disregarding the estimates obtained for the previous image pair [1,2]. However, several attempts have already been made to efficiently combine the temporal axis into the optical flow estimation process. Both Singh [3] and Chin & Willsky [4] proposed the application of Kalman filter as a mechanism to estimate the optical flow sequence in time. However, these proposed methods are computationally very complex. Fleet & Langley [5] proposed a different approach for the same task, based on Lucas & Kanade [2] optical flow method.

The purpose of this paper is to generalize Horn & Schunck algorithm to inherently include the temporal axis, while preserving simplicity and low computational algorithms. This paper is organized as follows: In Section 2 we present the differential framework and Horn & Schunck algorithm. Section 3 presents the new spatial-temporal optical flow model and recursive algorithms estimating it. Simulation results are presented in Section 4, and conclusion is in Section 5.

Note: A thorough description, analytical analysis and simulations can be found in [9].

2. Horn & Schunck Optical Flow estimation

The image sequence brightness is denoted by I(x,y,t), where (x,y,t) represent the spatial and temporal location. The brightness constraint equation is thus:

$$I(x,y,t) = I\left(x - dx(x,y,t), y - dy(x,y,t), t-1\right) \qquad (1)$$

where $\left[dx(x,y,t), dy(x,y,t)\right]$ is the motion vector corresponding to the pixel positioned at (x,y,t). Using Tailor series expansion, and neglecting higher derivative terms we get:

$$0 = dx(x,y,t)\frac{\partial I(x,y,t)}{\partial x} + dy(x,y,t)\frac{\partial I(x,y,t)}{\partial y} + \frac{\partial I(x,y,t)}{\partial t} \qquad (2)$$

The above equation poses one constraint per each pixel. Combining all the those equations together is possible by defining the following:

$$\underline{Y}(t) = -\left[I_t(1,1,t) \quad \cdots \quad I_t(x,y,t) \quad \cdots \quad I_t(N,N,t)\right]^T \qquad (3)$$

$$H(t) = \left[\mathrm{diag}\left\{I_x(x,y,t)\right\} \middle| \mathrm{diag}\left\{I_y(x,y,t)\right\}\right] \qquad (4)$$

$$\underline{X}(t) = \begin{bmatrix} \underline{D}_x \\ \underline{D}_y \end{bmatrix} \quad \begin{array}{l} \underline{D}_x^T = \left[dx(1,1,t) \quad \cdots \quad dx(N,N,t)\right] \\ \underline{D}_y^T = \left[dy(1,1,t) \quad \cdots \quad dy(N,N,t)\right] \end{array} \qquad (5)$$

Thus we have the model equation:

$$\underline{Y}(t) = H(t)\underline{X}(t) + \underline{E}(t) \qquad E\left\{\underline{E}(t)\underline{E}^T(t)\right\} = \sigma_e^2(t)I \qquad (6)$$

where $\underline{X}(t)$ is the optical flow to be estimated, and $\underline{E}(t)$ is the model error [1]. Additional spatial smoothness constraint should be combined in order to assure single solution and regularized problem. Denoting S as the Laplacian operator, the optical flow estimate should be the solution of the following quadratic minimization problem:

$$\hat{\underline{X}}(t) = \underset{\underline{X}(t)}{\arg\min}\left\{\left\|\underline{Y}(t) - H(t)\underline{X}(t)\right\|_2^2 + \beta\left\|\mathbf{S}\underline{X}(t)\right\|_2^2\right\} \quad (7)$$

The parameter β controls the relative smoothness required. The minimizing solution is:

$$\hat{\underline{X}}(t) = \left[H^T(t)H(t) + \beta\mathbf{S}^T\mathbf{S}\right]^{-1} H^T(t)\underline{Y}(t) \quad (8)$$

Instead of inverting the matrix shown above, the iterative Gauss-Siedel algorithm is suggested [1].

3. OF Estimation Along the Time Axis

Our aim is to propose a mechanism that will combine the time axis into the optical flow estimation process.

3.1 Kalman Filter (KF) approach

The Kalman filter the Minimum Mean Square Error (MMSE) estimate of the state of a linear system, represented by state-space equations [7]. Thus, in order to use the Kalman filter for the optical flow estimation task we must represent the problem in a state-space form. The unknown optical flow at time t, $\underline{X}(t)$, will serve as the state-vector to be estimated. The temporal smoothness constraint can be represented by:

$$\underline{X}(t) = \underline{X}(t-1) + \underline{N}(t), \quad E\left\{\underline{N}(t)\underline{N}^T(t-v)\right\} = W_N(t)\delta(v) \quad (9)$$

which simply says that the change in time in the optical flow is white (in time) vector $\underline{N}(t)$. Taking equation (6) and combining the regularization gives:

$$\begin{bmatrix}\underline{Y}(t)\\ \underline{0}\end{bmatrix} = \begin{bmatrix}H(t)\\ \mathbf{S}\end{bmatrix}\underline{X}(t) + \begin{bmatrix}\underline{E}(t)\\ \underline{F}(t)\end{bmatrix}$$
$$E\left\{\begin{bmatrix}\underline{E}(t)\\ \underline{F}(t)\end{bmatrix}\begin{bmatrix}\underline{E}(t-v)\\ \underline{F}(t-v)\end{bmatrix}^T\right\} = \begin{bmatrix}W_E(t) & 0\\ 0 & W_F(t)\end{bmatrix}\delta(v) \quad (10)$$

where the spatial smoothness serves here as an additional measurements of zeros. Having the above two equations enables the use of the Kalman filter directly. However, the dimensions of the matrices involved (though sparse) are very large, and direct application of the Kalman filter is impossible. In [4], a Square Root Information (SRI) Kalman filter is suggested [6] which propagates the square root of the inverse of the autocorrelation matrix in time. Yet, the computational complexity of the final algorithm is far too high, and only parallel implementation can cope with it effectively.

3.2 CWLS Approach

Instead of the state-space model presented above, we can suggest the following alternative model:

$$\forall k \ge 0 \quad \begin{bmatrix}\underline{Y}(t-k)\\ \underline{0}\end{bmatrix} = \begin{bmatrix}H(t-k)\\ \mathbf{S}\end{bmatrix}\underline{X}(t) + \underline{E}(t,k)$$
$$W_E(k-j) = \lambda^{-k}\begin{bmatrix}\sigma_e^2 I & 0\\ 0 & \sigma_f^2 I\end{bmatrix}\delta(k-j) \quad (11)$$

This model simply states that the optical flow vector $\underline{X}(t)$ matches the model equations for all casual times $t-k \le t$, and this way the temporal smoothness is applied. But, since we know that there are changes in the optical flow in time, we allow them by exponentially raising the variance of the model error for far away model equations, and the parameter $0 << \lambda < 1$ acts as a forgetting factor for this very purpose. Having the new model, we can define a quadratic error:

$$\varepsilon^2(t) = \sum_{k=0}^{\infty}\lambda^k\left\{\left\|\underline{Y}(t-k) - H(t-k)\underline{X}(t)\right\|_2^2 + \beta\left\|\mathbf{S}\underline{X}(t)\right\|_2^2\right\} \quad (12)$$

Differentiating with respect to the vector $\underline{X}(t)$ yields the following equations:

$$\frac{\partial\varepsilon^2(t)}{\partial\underline{X}(t)} = 0 = 2\left[R(t)\underline{X}(t) - \underline{P}(t)\right] \quad (13)$$

$$R(t) = \lambda R(t-1) + H^T(t)H(t) + \beta\mathbf{S}^T\mathbf{S} \quad (14)$$

$$\underline{P}(t) = \lambda\underline{P}(t-1) + H^T(t)\underline{Y}(t) \quad (15)$$

3.3 The P-RLS OF Estimation Algorithm

One way to solve the minimization problem in equation (12) is a direct solution of the linear system in equation (13). The matrix R(t) is a positive - thus ensuring unique solution. The matrix R(t) is also sparse and can be easily updated in time and stored. The number of unknowns in the vector $\underline{X}(t)$ is $2N^2$ and the size of the matrix R(t) is $[2N^2 \times 2N^2]$. Since N is typically large, this means that a direct inversion of R(t) is impossible and indirect methods are required in order to solve (13). Many iterative algorithms can be suggested [8]. The reason we call such procedures Pseudo-RLS algorithms comes from the fact that we update the matrix R(t) and the vector $\underline{P}(t)$ recursively, as can be done in the Recursive Least Squares (RLS) algorithm [6]. However, in contrast to the classic RLS, we do not propagate nor compute the matrix $Q(t) = R^{-1}(t)$.

The amount of computations required is similar to the amount required by the original Horn & Schunck algorithm. This is because we need to compute the update terms for the matrix R(t) and the vector $\underline{P}(t)$, which are exactly the terms computed for the Horn & Schunck algorithm, we have to add them to R(t) and $\underline{P}(t)$ which require only additions, and then we have to apply an iterative algorithm similar to what is done in the original Horn & Schunck algorithm.

One important question is the connection between the KF and the proposed estimation approach. In [9], a thorough analysis is given, showing that the Pseudo-RLS algorithm yields an unbiased, and bounded variance estimation, compared to the KF.

3.4 The M-SD and the M-LMS Algorithms

A different approach that can be taken in order to minimize the temporal squared error in equation (12) is suggested by the Least Mean Square (LMS) algorithm [6]. First, instead of a full minimization of this error at each time instant, we can simply take the previous result $\hat{\underline{X}}(t-1)$ and update it using the instantaneous gradient of the temporal squared error and get the following recursive equation:

$$\begin{aligned}\hat{\underline{X}}(t) &= \hat{\underline{X}}(t-1) - \frac{\mu}{2}\frac{\partial \varepsilon^2(t)}{\partial \underline{X}(t)}\bigg|_{\hat{\underline{X}}(t-1)} = \\ &= \hat{\underline{X}}(t-1) + \mu\left[\underline{P}(t) - R(t)\hat{\underline{X}}(t-1)\right]\end{aligned} \tag{16}$$

Equation (16) is simply one iteration of the Steepest Decent algorithm. Thus, instead of performing many iterations at each time instant as was proposed in the Pseudo-RLS algorithm, all we are proposing to do here is a single iteration, and continue to the next temporal point. Therefore, we can suggest also a midway algorithm, namely, at each temporal point, update the matrix and the vector R(t) and $\underline{P}(t)$ as usual using equations (14) and (15), and then perform M Steepest Decent iterations. We already know that for $1 \ll M \to \infty$ we get the Pseudo-RLS. We will refer to the above algorithm with M iterations per each time instant as M-SD algorithm. A desired property of the M-SD algorithm is it's flexibility with regard to the computational requirements. The more iterations performed the better is the quality of the estimated optical flow. Moreover, the M-SD estimation error with fixed M is bounded [9].

Using equation (16) as our estimation process is an approximation of the Pseudo-RLS estimator presented in equation (12). Using the recursive equations (14) and (15) in equation (16), we get:

$$\begin{aligned}\hat{\underline{X}}(t) &= \hat{\underline{X}}(t-1) + \mu\lambda\left[\underline{P}(t-1) - R(t-1)\hat{\underline{X}}(t-1)\right] + \\ &+ \mu H^T(t)\underline{Y}(t) - \mu\left[H^T(t)H(t) + \beta \mathbf{S}^T\mathbf{S}\right]\hat{\underline{X}}(t-1)\end{aligned} \tag{17}$$

If we assume that the previous solution $\hat{\underline{X}}(t-1)$ is close to the optimal one, we can say that $\underline{P}(t-1) - R(t-1)\hat{\underline{X}}(t-1) \approx 0$, and this term can be omitted from the above equation, yielding

$$\begin{aligned}\hat{\underline{X}}(t) &= \hat{\underline{X}}(t-1) + \mu H^T(t)\underline{Y}(t) - \\ &- \mu\left[H^T(t)H(t) + \beta \mathbf{S}^T\mathbf{S}\right]\hat{\underline{X}}(t-1)\end{aligned} \tag{18}$$

which is a simpler algorithm with even more reduced computations, since this algorithm no longer requires the propagation of R(t) and $\underline{P}(t)$ in time. The above equation is a single SD iteration of the original Horn & Schunck algorithm. Interesting as it may seem, Horn & Schunck in their original paper [1] suggested this very algorithm based on intuition only as an alternative to the application of their algorithm with many iterations per one step. Following the same reasoning as in the M-SD algorithm, we can propose here that M iterations of equation (18) can be performed per one time step, which might improve the overall performance of the algorithm, when compared to single iteration algorithm. We choose to call such algorithm the M-LMS algorithm for obvious reasons.

4. Simulations and Analysis

The presented test is performed on a synthetic image sequence with a-priori known optical flow, in order to quantify the results. The tested image sequence consist of a constant rotation of 1.2° per image with additional zoom in and out in the form of half a cycle of a sine function in the range [0.85-1.15]. The optical flow sequence thus change in time. The test sequence contains 101 images of size $[50 \times 50]$ pixels. A Gaussian white random noise with variance $\sigma_n = 4$ is added to each image, where the dynamic range of the images is [0,255].

Figure 1 presents the actual optical flow, the estimation results using the M-SD, and the M-LMS algorithms. The M-SD and the M-LMS results were obtained using M=5 iterations, whereas the Horn & Schunck results were obtained by performing 200 NSD iterations per each temporal point, thus using much more computations. Based on these results, we can say that the performance of the M-SD and the M-LMS algorithms on this test are significantly better than the Horn & Schunck approach.

5. Conclusion

In this paper we have presented new algorithms for the estimation of optical flow for image sequences. These new algorithms are based on Horn & Schunck algorithm, generalized to include temporal smoothness. The undertaken approach starts from state-space equations modeling the estimation problem, but instead of applying Kalman filter which seems natural at this point, we further simplify the model leading to adaptive filtering formulations resembling the RLS and the LMS algorithms [6]. The new estimation methods are shown to give low complexity requirement, while providing more accurate results compared to Horn & Schunck algorithm.

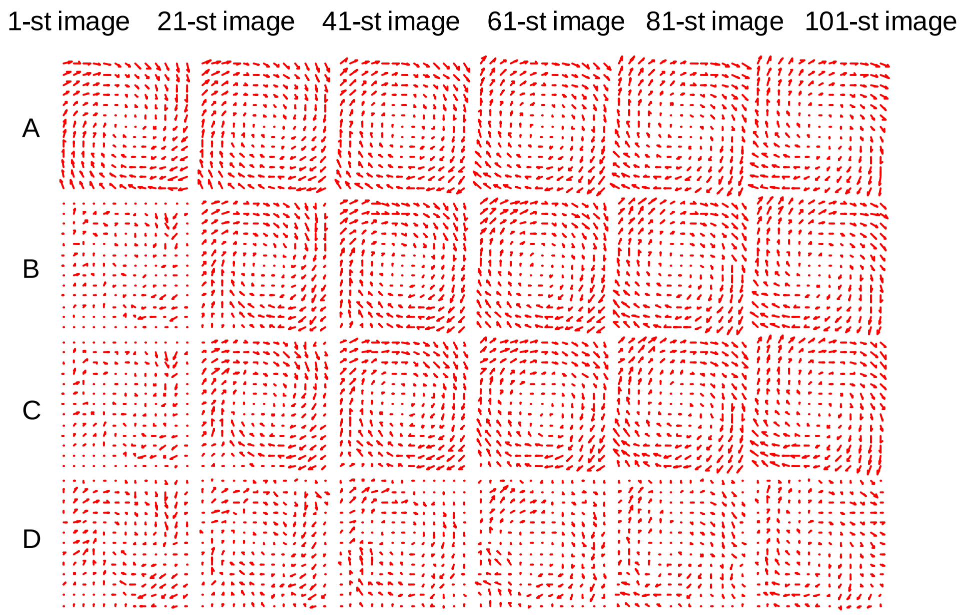

Figure 1 - The Optical Flow A: true, B: M-SD, C: M-LMS, D: Horn & Schunck

References

[1] B. K. P. Horn and B. G. Schunck, "Determining Optical Flow", AI, Vol. 17, pp. 185-204, 1981.

[2] B. Lucas & T. Kanade, "An Iterative Image Registration Technique with application to Stereo Vision", Proc. DARPA Image Understanding Workshop, pp. 121-130, 1981.

[3] A. Singh, "Incremental Estimation of Image Flow Using the Kalman Filter", J. Visual Communication and Image Representation, Vol. 3, pp. 39-57, 1992.

[4] T. M. Chin, W. C. Karl, A. M. Mariano & A. S. Willsky, "Square Root Filtering in Time-Sequential Estimation of Random Fields", Proc. SPIE, Vol. 1903, pp. 51-58, 1993.

[5] D. J. Fleet & K. Langley, "Recursive Filters for Optical Flow", IEEE Trans. Pattern analysis and Machine Intelligence, Vol. 17, pp. 61-67, 1995.

[6] S. Haykin, Adaptive Filter Theory, Prentice-Hall, 1986.

[7] C. K. Chui & G. Chen, Kalman Filtering, Springer-Verlag, 1990.

[8] L. A. Hageman & D. M. Young, Applied Iterative Methods, Academic Press, 1981.

[9] M. Elad & A. Feuer, “Recursive Optical Flow Estimation - Adaptive Filtering Approach”, Internal Report #1009, the Technion - Israel Institute of Technology, 1995.