

Pel-Recursive Motion Estimation Algorithm

By

H. Gharavi
National Institute of Standards & Technology (NIST)
100 Bureau DR Stop8920, Gaithersburg, MD 20899-8920, USA.

And

H. Reza-Alikhani
Loughborough University
Loughborough, Leics., LE11 3TU, UK.

Abstract: This paper presents a new pel recursive motion compensated prediction algorithm for video coding applications. The derivation of the algorithm is based on Recursive Least Squares (RLS) estimation that minimizes the mean square prediction error for each pel (picture element). A comparison with the modified Steepest-descent gradient estimation algorithm shows significant improvement in terms of mean-square prediction error performance.

Introduction: Netravali and Robbins [1] developed a pel recursive spatio-temporal steepest-descent gradient technique in which the displacement of a pel was predicted from previously transmitted information. Since then various algorithms have been proposed to improve the performance of pel recursive motion estimation (PRME) techniques. The most important contribution was the modification of the steepest-descent algorithm which was introduced by Walker and Rao [2]. In this paper we present a simple but very efficient PRME algorithms that can significantly outperform the modified steepest-descent technique developed by Walker and Rao.

Proposed Algorithm: For the sake of our analysis, we assume the translation movement of an object is in a plane parallel to the camera and illumination is uniform. We also assume the effect of uncovered background to be negligible. Under these assumptions, let $S(x, y, t)$ denote the monochrome intensities at point x, y of a moving object in the image plane where its transnational

movement is at the constant velocity of v_x and v_y . We can show that after Δt second (one frame period), the object moves to a new location where we can show,

$$S(x, y: t + \Delta t) = S[(x + v_x \Delta t), (y + v_y \Delta t): t] \quad (1)$$

Expanding the field in a power series in Δt and neglecting the higher order terms,

$$S(x, y: t + \Delta t) = S(x, y: t) + \frac{\partial}{\partial x} S(x, y: t) [v_x] \Delta t + \frac{\partial}{\partial y} S(x, y: t) [v_y] \Delta t \quad (2)$$

or

$$S(x, y: t + \Delta t) - S(x, y: t) = \frac{\partial}{\partial x} S(x, y: t) dx + \frac{\partial}{\partial y} S(x, y: t) dy \quad (3)$$

where d_x and d_y correspond to the horizontal and vertical components of the motion displacement vector D . Assuming that $\frac{\partial}{\partial x} S(x, y: t)$ and $\frac{\partial}{\partial y} S(x, y: t)$ are known for each x, y, t and defining $ED, LD,$ and FD as the magnitude of the element, line, and frame difference at point n , from (3), we can write,

$$FD = \Phi_n^T D \quad (4)$$

Where $()^T$ is transpose and,

$$\Phi_n = \begin{bmatrix} \frac{\partial}{\partial x} S(x_n, y_n: t) \\ \frac{\partial}{\partial y} S(x_n, y_n: t) \end{bmatrix} = \begin{bmatrix} ED \\ LD \end{bmatrix} \quad (5)$$

From (4) the frame difference (FD) measurement can be shown as,

$$\xi_n = \Phi_n^T \bar{D} + \text{noise} \quad (6)$$

where $\bar{D} = [\bar{d}(x), \bar{d}(y)]^T$ is the motion vector estimate.

For a cluster of M moving pels, the least-squares estimate of D can be obtained as,

$$LSE = \min_D \left\{ \sum_{n=1}^m (\xi_n - \Phi_n^T \bar{D})^2 \right\} \quad (7)$$

After minimization,

$$\sum_{n=1}^m \Phi_n \xi_n = \bar{D} \sum_{n=1}^m \Phi_n \Phi_n^T \quad (8)$$

For,

$$\eta = \frac{1}{M} \sum_{n=1}^M \Phi_n \xi_n \quad \text{and} \quad R = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T \quad (9)$$

the estimated motion vector from (8) is obtained as,

$$\bar{D} = R^{-1} \eta \quad (10)$$

$$\begin{aligned} \eta_i &= \eta_{i-1} + \Phi_n \xi_n \\ R_i &= R_{i-1} + \Phi_n \Phi_n^T \end{aligned} \quad (11)$$

For recursive estimation of η and R , we can write

Based on the so-called matrix inversion lemma, the inverse of R_i can be obtained as,

From (10), (11), and (12),

$$R_i^{-1} = R_{i-1}^{-1} - \frac{R_{i-1}^{-1} \Phi_n \Phi_n^T R_{i-1}^{-1}}{1 + \Phi_n^T R_{i-1}^{-1} \Phi_n} \quad (12)$$

$$\bar{D}_i = \bar{D}_{i-1} - \frac{A_{i-1}^{-1} \Phi_n}{1 + \Phi_n^T A_{i-1}^{-1} \Phi_n} (\Phi_n^T \bar{D}_{i-1} - \xi) \quad (13)$$

In the above equation, The term in the right hand side bracket can be replaced by what is known as the Displaced Frame Difference, DFD. Thus,

$$\bar{D}_i = \bar{D}_{i-1} - \frac{A_{i-1}^{-1} \Phi_n}{1 + \Phi_n^T A_{i-1}^{-1} \Phi_n} [DFD(x, y, \bar{D}_{i-1})] \quad (14)$$

Simplifications:

To avoid matrix inversion at each iteration, (14) can be simplified by neglecting the x and y cross terms in calculating ϕ_n and R . Thus, from (5) and (9),

$$F_n(x) = ED \quad \text{and} \quad F_n(y) = LD \quad (15)$$

$$R(x) = \frac{1}{M} \sum_j^M ED_j^2 \quad \text{and} \quad R(y) = \frac{1}{M} \sum_j^M LD_j^2$$

Applying (15) to (14), the components of the motion displacement estimates can be shown as,

$$\begin{aligned}\bar{d}_i(x) &= \bar{d}_{i-1}(x) - \frac{ED}{\sum ED^2 + ED^2} \{DFD[x, y, \bar{d}_{i-1}(x)]\} \\ \bar{d}_i(y) &= \bar{d}_{i-1}(y) - \frac{LD}{\sum LD^2 + LD^2} \{DFD[x, y, \bar{d}_{i-1}(x)]\}\end{aligned}\tag{16}$$

Simulation Results

The computation involved in (16) is performed recursively. At each iteration the estimated motion displacement is applied to measure a new DFD. This would first require obtaining the location of the displaced pel on the previous frame, based on the estimated components of motion displacement. Since the motion estimates are expected to be non-integer, the luminance value of the displaced pel is predicted by a two dimensional interpolator which uses the four corners of the surrounding pels in a two dimensional grid. In our experiments, the DFD is measured at two locations with reference to the current pel; the pel above (i.e., previous line), and the previous pel along the same line. The average of the two DFD's (with equal weightings) is then used to update the displacement estimates.

The ED and LD in (16) were also measured using the interpolated luminance values from the displaced previous frame. For $\sum ED^2$ and $\sum LD^2$ the summation includes the luminance values of three interpolated neighboring pels from the previous frame. The above algorithm is applied to a cluster of pels that are classified as moving areas. In our experiments, a moving area consists of three neighboring pels (including the current pel) whose frame difference exceeds a predefined threshold (i.e. $|FD| > \text{threshold}$).

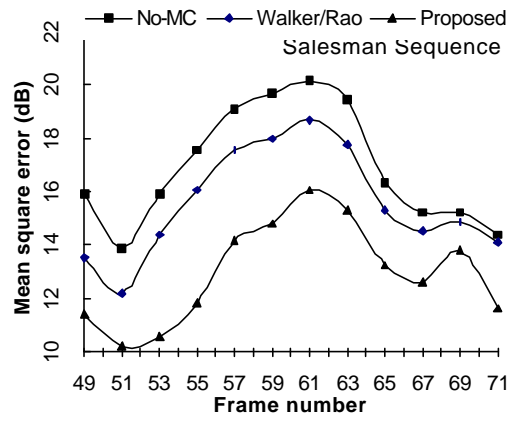
Two video sequences, known as "Salesman" and "Suzie," were considered for evaluating the performance of the proposed algorithm. The format of both sequences was based on the CIF (Common Intermediate Format: 352-pels by 288-lines and 30 frames/s). In addition, for the sake of comparison, we have simulated the Walker-Rao algorithm [2]. The simulation results of both schemes, in terms of mean square prediction error (in dB), are shown in Figures 1 and 2 for the

“Salesman” and “Suzie” sequences respectively. In these figures we have also included the results of interframe prediction without motion compensation (i.e., frame difference). The number of iterations for both schemes was 3 and the threshold value was 9. In addition, these results were obtained using the second previous frame for prediction (i.e., skipping one frame). Looking at these figures, it can be clearly observed that the proposed scheme can significantly reduce the motion compensated prediction error. In terms of subjective evaluations, Figure 3 presents the motion compensated prediction error images between frames 49 and 51 of the “Suzie” sequence. In these images, relatively darker or lighter patches represent the degree of inaccuracies in estimating the components of the motion displacement. Comparing the two images confirms the performance superiority of the proposed scheme over the modified steepest-descent algorithm, particularly in the regions where the motion activities are relatively high.

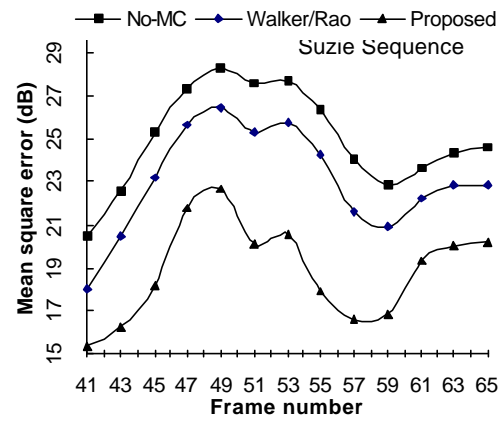
Conclusion: This paper proposes an efficient pel-recursive estimation technique for motion tracking and coding of moving images. The proposed algorithm has been compared with the modified steepest-descent gradient algorithm. The results indicate a considerable reduction in the prediction error, particularly in regions where the motion activities are relatively high.

References

- [1] A. N. Netravali and J. D. Robbins, "Motion compensated television coding- part I", Bell Syst. Tech. J., Vol. 58, pp. 631-670, Mar. 1979.
- [2] D.R. Walker and K.r. Rao, “Improved Pel-recursive motion-estimation”, IEEE Trans. Commun, Vol. COM-32, No. 10, 1984, pp. 1128-1134



(a)

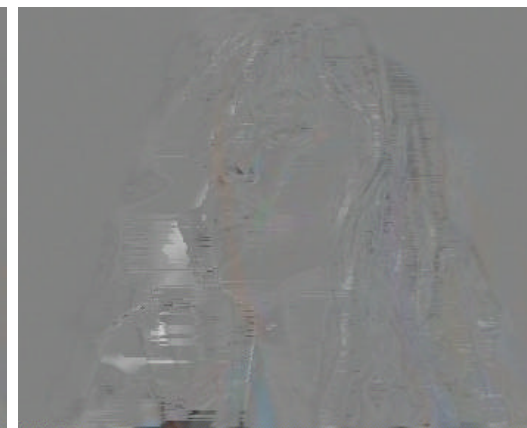


(b)

Figure 1: Mean-square error performance using the second previous frame for prediction (a) Salesman sequence, (b) Suzie sequence



(a)



(b)

Figure 2: Motion compensated prediction error images for Suzie sequence (a): Walker & Rao scheme (b): Proposed scheme.