MUSMANN

A Differential Displacement Estimation Algorithm With Improved Stability

M. Bierling
Institut für Theoretische Nachrichtentechnik
und Informationsverarbeitung
Universität Hannover
Callinstr. 32
3 Hannover 1
F.R. Germany

Abstract

A new differential displacement estimation algorithm for television sequences is presented. It minimizes the local mean squared displaced frame difference rather than maximizing the local cross correlation of displaced frames, as it can be shown that there is frequently no correspondence between the cross correlation peak and the actual displacement of a moving object. The algorithm is applied iteratively, i.e. in each step of iteration the resulting estimate of the displacement vector serves as an initial guess for the next step. Compared to known techniques stability is improved by introducing a more accurate two-dimensional image model. The approximation of spatial gradients as an average of spatial differences of two successive frames yields an increased accuracy of the displacement estimate.

Introduction

Temporal luminance changes of successive pictures in television sequences are oftentimes due to the motion of objects. The velocity parameters obtained by a motion estimation algorithm can be applied to motion compensation techniques, such as motion compensated predictive coding, frame interpolation, or noise reduction. Huang [1] has proposed an approach to estimate the motion parameters for the model of three-dimensional rigid objects with translatory and rotatory motion. With respect to the requirements of real-time digital signal processing the complexity of a motion estimation algorithm has to be relatively small. Hence, several methods have been derived, which estimate a displacement vector assuming objects translatorily displaced in the image plane. Two main classes of algorithms are known, matching techniques and spatio-temporal gradient methods [2].

Frequently the measurement of translatory displacement is performed by determining the peak of the local cross correlation function of displaced frames [3],[4]. In order to estimate the local maximum of the cross correlation function, e.g. a gradient algorithm is applied for rectangular measurement windows containing small image regions of two successive frames. This approach fails, if there is no correspondence between the cross correlation peak and the actual displacement of a moving object. This happens for example in the case of a translatorily displaced luminance plane, which corresponds to a slight luminance edge. Then the normalized cross correlation is constant and the displacement estimate is mainly affected by the present noise. Furthermore, investigations of Beyer [5] have shown that an evaluation of the luminance signal on small measurement windows yields an unsymmetrical cross correlation function and therefore an estimation error is inherent, even if the correlation peak corresponds with the actual displacement. Other known displacement estimation algorithms minimize the local mean squared displaced frame difference rather than maximizing the local cross correlation function of displaced frames [5]-[8]. In the case of pure translatory displacement the displaced frame difference always has a local minimum corresponding with the actual displacement.

In this contribution a differential displacement estimation algorithm is presented, that minimizes the local mean squared displaced frame difference by means of a spatio-temporal gradient method. It is related to the algorithms proposed by Cafforio and Rocca [6] and Bergmann [7]. Stability is improved using a more accurate two-dimensional image model. First the algorithm is derived, then experimental results are presented.

The algorithm

Assume a pure translatorily moving object that does not change its luminance from frame k-1 to frame k. The luminance $S_{k-1}(x,y)$ at the position x,y in frame k-1 is given by

$$S_{k-1}(x,y) = S_k(x+dx,y+dy) \tag{1}$$

where dx,dy are the components of the displacement vector $\vec{D}$. Using a Taylor series expansion to express the luminance function we obtain

$$
\begin{aligned}
S_k(x+dx,y+dy) = \; & S_k(x,y) \\
& + \{\partial S_k(x,y)/\partial x\}\cdot dx + \{\partial S_k(x,y)/\partial y\}\cdot dy \\
& + \{\partial^2 S_k(x,y)/\partial x^2\}\cdot dx^2/2 \\
& + \{\partial^2 S_k(x,y)/\partial y^2\}\cdot dy^2/2 \\
& + \{\partial^2 S_k(x,y)/\partial x \partial y\}\cdot dxdy + r(x,y)
\end{aligned} \tag{2}
$$

where r(x,y) denotes the higher order terms of the taylor series expansion, which will be neglected in the following. In contrast to the algorithm in [6], the second order derivatives are not discarded. This image model is more accurate than a simple linear function. Assuming r(x,y) to be zero we obtain with (1) and (2)

$$S_k(x+dx,y+dy) = S_k(x,y) + \bar{G}_x(x,y)\cdot dx + \bar{G}_y(x,y)\cdot dy$$

with

$$
\begin{aligned}
\bar{G}_x(x,y) &= [\partial S_k(x,y)/\partial x + \partial S_{k-1}(x,y)/\partial x] \;/\; 2 \\
\bar{G}_y(x,y) &= [\partial S_k(x,y)/\partial y + \partial S_{k-1}(x,y)/\partial y] \;/\; 2
\end{aligned} \tag{3}
$$

Thus the second order derivatives are taken into account by averaging the spatial gradients of two successive frames. The frame difference $FD(x,y,\vec{D})$ generated by a displacement $\vec{D}$ results as

$$
\begin{aligned}
FD(x,y,\vec{D}) &= S_k(x,y) - S_{k-1}(x,y) \\
&\approx -\bar{G}_x\cdot \hat{dx} - \bar{G}_y\cdot \hat{dy} \\
&= \hat{FD}(x,y,\hat{\vec{D}})
\end{aligned} \tag{4}
$$

With this approximation the displacement vector $\vec{D}$ can be estimated by $\hat{\vec{D}} = (\hat{dx},\hat{dy})^T$ such that the mean squared difference between the measured frame difference FD and the approximated frame difference $\hat{FD}$ is minimized

$$E[\{ FD(x,y,\vec{D}) - \hat{FD}(x,y,\hat{\vec{D}}) \}^2] \rightarrow Min. \tag{5}$$

which is exactly the same as minimizing the mean squared displaced frame difference as a function of the estimated displacement vector

$$DFD(x,y,\hat{\vec{D}}) = S_k(x+\hat{dx},y+\hat{dy}) - S_{k-1}(x,y) \tag{6}$$

The solution is obtained by setting the partial derivatives with respect to the vector components $\hat{dx},\hat{dy}$ of the left hand side of (5) to zero and solving a system of two linear equations, which yields

$$
\begin{aligned}
\hat{dx} &= \{ E[\bar{G}_x\cdot\bar{G}_y]\cdot E[FD\cdot\bar{G}_y] - E[FD\cdot\bar{G}_x]\cdot E[\bar{G}_y^2] \} \;/\; DET \\
\hat{dy} &= \{ E[\bar{G}_x\cdot\bar{G}_y]\cdot E[FD\cdot\bar{G}_x] - E[FD\cdot\bar{G}_y]\cdot E[\bar{G}_x^2] \} \;/\; DET
\end{aligned}
$$

with (7)

$$DET = E[\bar{G}_x^2]\cdot E[\bar{G}_y^2] - E^2[\bar{G}_x\cdot\bar{G}_y]$$

where the coordinates x,y are omitted for simplicity. Of course, in digital video processing the expected values have to be approximated by summing over measurement windows, placed in the moving area of two successive frames. An estimate obtained from equation (7) is then assigned to the center of a measurement window. The spatial gradients of the luminance signal can be calculated by centered differences, adopting a proposal of Cafforio and Rocca [8]. The denominator DET in equation (7) is the determinant of the set of linear equations. Therefore no unique solution is obtainable, if the denominator is zero. In fact, using relatively small measurement windows the displacement vector is ambiguous if the signal gradient locally is zero in one arbitrary direction. It is easy to show that in this case the ratio $\overline{G}_x/\overline{G}_y$ is a constant value and therefore the denominator DET in (7) is zero. To overcome this problem, the addition of a positive constant in the denominator has been suggested by Cafforio and Rocca [8]. Nevertheless, it is possible to obtain a more reasonable solution in the case of a locally ambiguous displacement vector by selecting the vector with minimum magnitude, minimizing the local displaced frame difference at the same time. With the additional constraint of a minimum magnitude vector and with zero denominator DET we obtain instead of equation (7)

$$\hat{d}x = - E[FD \cdot \overline{G}_x] \; / \; \{ E[\overline{G}_x^2] + E[\overline{G}_y^2] \}$$
$$\hat{d}y = - E[FD \cdot \overline{G}_y] \; / \; \{ E[\overline{G}_x^2] + E[\overline{G}_y^2] \} \tag{8}$$

The displacement estimate can be improved by a motion compensated iteration of the algorithm, i.e. in the i-th step of iteration the coordinates x,y of all terms belonging to frame k-1 are substituted by $x-\hat{d}x_{i-1}, y-\hat{d}y_{i-1}$ to obtain a more accurate estimate $\hat{D}_i$. In the i-th iteration, for example

$$S_{k-1,i}(x,y) = S_{k-1,i-1}(x-\hat{d}x_{i-1}, y-\hat{d}y_{i-1}) \tag{9}$$

Additionally, a spatial recursion can increase the accuracy of the estimate by using the estimate previously calculated for adjacent picture elements, as has been pointed out in [8].

The estimation algorithm given by equation (7) is related to the algorithm proposed by Cafforio and Rocca [6], however, the accuracy of the estimate has been increased by introducing a two-dimensional quadratic image model. The algorithm by Cafforio and Rocca assumes a linear luminance function. Furthermore it should be noted that the consideration of the correlation of the components of the spatial gradient vector is important to ensure a sufficient stability behaviour. By neglecting the terms $E[\overline{G}_x \cdot \overline{G}_y]$ in equation (7) and with the additional assumption of stationary luminance signal we obtain

$$\hat{d}x = - E[FD \cdot \overline{G}_x] \; / \; E[\overline{G}_x \cdot \partial S_k(x,y)/\partial x]$$
$$\hat{d}y = - E[FD \cdot \overline{G}_y] \; / \; E[\overline{G}_y \cdot \partial S_k(x,y)/\partial y] \tag{10}$$

which is nearly identical to the algorithm proposed by Bergmann [7]. The assumption of a zero correlation of the components of the spatial gradient vector is only valid for large measurement windows and obviously yields an algorithm based on an one-dimensional image model. This results in an insufficient stability especially at moving diagonal structures, as shown by experimental evaluations. On the other hand the complexity of the algorithm given by equation (10) is only little smaller than that one given by equation (7), considering that the approximation of the expected values is the main computational load: equation (10) requires the computation of 4, equation (7) of 5 distinct expected values.

Experimental results

The displacement estimation algorithm given by equation (7) has been compared experimentally to the algorithms of Cafforio and Rocca [6] and Bergmann [7]. The test sequence, recorded with a camera, sampled at 10 MHz and quantized according to 8 bit/sample, shows vertical and diagonal black bars moving at a horizontal velocity of about 5.4 pel/frame. The evaluations have been carried out only with the odd fields of two successive frames to avoid effects from line interleaving. Fig. 1a shows the first of two fields, which have been used. For each picture element the displacement vector has been estimated independently, i.e. no spatial recursion has been applied. The measurement windows consisted of 13 by 13 pels, and 3 steps of iteration were used. In order to achieve an estimate with sub-pel accuracy the luminance values according to equation (9) were interpolated in the case of nonintegral displacement vector components. Fig. 1b shows the frame differences caused by the motion, where positive and negative differences are displayed with white and black, zero differences with gray luminance values. Fig. 1c - 1e show the displaced frame differences obtained by applying the considered estimation algorithms. Table 1 shows the entropies and variances of the displaced frame differences for the different algorithms. The algorithm of Bergmann only performs well in those image parts, where the assumption of an one-dimensional image model is valid. The algorithm given by equation (7) proves most favorable as a result of a more accurate image model. This is confirmed by the variances of the displaced frame differences listed in Table 1.

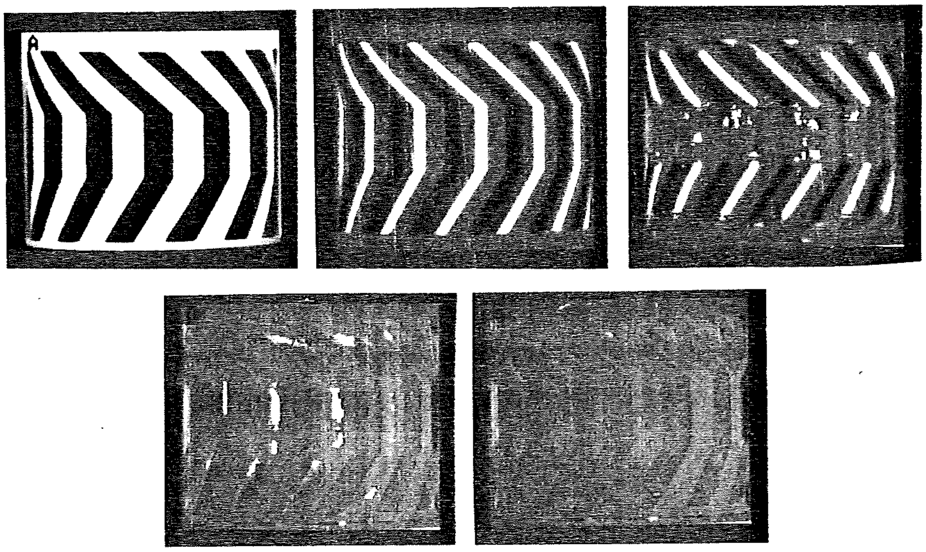

Figure 1 Illustration of the displaced frame differences obtained with several estimation algorithms. One field of the test sequence used for the computer simulations is shown in a) (<- : direction of movement). Displaced frame differences obtained b) without motion compensation, c) using the algorithm of Bergmann [7], d) using the algorithm of Cafforio and Rocca [6], e) using the algorithm given by equation (7).

a b c
d e

Table 1 Entropies and variances of the displaced frame differences shown in Figure 1

Algorithm	Entropy	Variance
no motion compensation	6.2 bit/pel	2921
Bergmann [7]	5.8 bit/pel	1368
Cafforio and Rocca [6]	4.4 bit/pel	148
Equation (7)	4.2 bit/pel	37

Conclusion

A new differential displacement estimation algorithm for digital television sequences has been presented. It minimizes the local mean squared displaced frame difference. The algorithm is applied iteratively, i.e. in each step of iteration the resulting estimate of the displacement vector serves as an initial guess for the next step. Compared to known techniques stability behaviour and the precision of the estimate is improved by a two-dimensional quadratic image model. Experimental results confirm that the algorithm performs well, even for small measurement windows.

Further investigations should extend the convergency range of the algorithm by prefiltering the picture signal, applied in the first steps of iteration to cope with large initial displacements.

References

[1] T.S. Huang and R.Y. Tsai, "Image sequence analysis: Motion Estimation," in **Image Sequence Analysis**. Berlin, Germany: Springer-Verlag, 1981, pp. 1-18.

[2] H.G. Musmann, P. Pirsch, and H.-J. Grallert, "Advances in Picture Coding," Proc. IEEE, vol. 73, pp. 523-530, Apr. 1985.

[3] H.C. Bergmann, "Analysis of Different Displacement Estimation Algorithms for Digital Television Signals," in **Image Sequence Processing and Dynamic Scene Analysis**, edited by T.S. Huang, Berlin, Germany: Springer-Verlag, pp. 215-234, 1983.

[4] S. Beyer, "Maximum Likelihood Estimation of Translatoric Displacement in Image Sequences," NTG-Fachberichte, Bd. 84, Berlin, Germany: VDE-Verlag, pp. 327-339, 1983.

[5] S. Beyer, "Displacement-Schätzverfahren für Fernsehbildsignale mit minimaler Schätzfehlervarianz," Ph.D. dissertation, Univ. of Hannover, Hannover, Germany, 1985.

[6] C. Cafforio and F. Rocca, "Methods for measuring small displacements of television images," IEEE Trans. Inform. Theory, vol. IT-22, no. 5, pp. 573-579, Sept. 1976.

[7] H.C. Bergmann, "Ein schnell konvergierendes Displacement-Schätzverfahren für die Interpolation von Fernsehbildsequenzen," Ph.D. dissertation, Univ. of Hannover, Hannover, Germany, Feb. 1984.

[8] C. Cafforio and F. Rocca, "The differential method for image motion estimation," in **Image Sequence Processing and Dynamic Scene Analysis**, edited by T.S. Huang, Berlin, Germany: Springer-Verlag, pp. 104-124, 1983.