Noise Reduction in Image Sequences with Sparse Temporal Sampling

Dirk Borghys and Marc Acheroy

Royal Military Academy, Electrical Engineering Department,
30, Av. de la Renaissance, 1040 Bruxelles, BELGIUM

Tel: 32-2-737.61.61, Fax: 32-2-737.61.63, Email: dirk@elec.rma.ac.be, http://www.elec.rma.ac.be/~dirk

Abstract— **In this paper, a method for noise reduction in image sequences with sparse temporal sampling is proposed. The method combines motion compensation with an adaptive temporal low-pass filter. Image sequences are mostly temporally non-stationary due to sensor motion, the motion of individual objects in the scene and objects which disappear, (re)appear or chang their orientation. The motion component can be compensated for by estimating the optic flow through the sequence. This allows to greatly reduce the non-stationarities. However, not all non-stationarities can be accounted for by motion. Therefore, after motion compensation an adaptive filter is used to perform the actual noise reduction.**

Keywords— **restoration, noise reduction, motion compensation, image sequence**

I. Introduction

Image sequences are used in a wide variety of applications: video communications, target detection and tracking, object recognition, medical imaging, etc... These sequences are mostly corrupted by random noise at various stages. (generation, transmission or recording) Hence the importance of reducing this noise.

A temporal low-pass filter is useful to reduce the noise in the image sequence but it will blur the non-stationary regions in the image. To preserve the moving regions during temporal smoothing, motion compensation is used. For this purpose, the optic flow field is estimated between each pair of subsequent images of the sequence. For the estimation of optic flow a multi-resolution method is used which is based on conservation constraints and a smoothness constraint. For the conservation constraints, a weighted sum of several image properties is used.

Then, the images are warped using this flow field. After the warping, a temporal low-pass filter would reduce noise without blurring the moving regions in the images. However, non-stationarities that are not due to motion will still be blurred. Therefore an adaptive version of the temporal low-pass filter is used.

II. Motion Compensation

A. optic Flow Estimation

The method used for estimating the optic flow between two images of a sequence is based on [1] and is divided into two parts. The first part is based on the conservation of certain local measures. The second part exploits the continuity of the flow field using so called neighbourhood constraints. These constraints state that a pixel in an image is more likely to move in a similar fashion as its neighbours.

A.1 Conservation constraint

Let P_k be the k^{th} property that is conserved. A rectangle of dimension $w \times h$ is defined around each pixel in the first image and we wish to know the position in the second image that gives the most similar value for the different properties. We will therefore displace the rectangle (by (u, v)) and each time calculate an error function. This error function is defined as:

$$\begin{aligned} E_{xy}(u,v) &= \sum_k \gamma_k \sum_{i=-\frac{w}{2}}^{\frac{w}{2}} \sum_{j=-\frac{h}{2}}^{\frac{h}{2}} \left[P_k^{im1}(x+i, y+j)\right. \\ &\quad \left. - P_k^{im2}(x+i+u, y+j+v)\right]^2 \end{aligned} \tag{1}$$

γ_k allows to put different weights on different conserved properties. The following properties were used:

- The grey value.
- The norm of the gradient
- The direction of the gradient.
- Texture measures (Haralick parameters [2]).

Using this definition of the error function, a response function is defined as:

$$R_c(u,v) = exp[-\kappa E_{xy}(u,v)] \tag{2}$$

This response function can be seen as the probability density function of the velocity distribution due to the conservation constraint. The displacement which corresponds best to the conservation constraint at a given point (x, y) in the image can then be defined as the weighted average:

$$u_c(x,y) = \frac{\sum_u \sum_v R_c(u,v)u}{\sum_u \sum_v R_c(u,v)} \quad (3)$$

$$v_c(x,y) = \frac{\sum_u \sum_v R_c(u,v)v}{\sum_u \sum_v R_c(u,v)} \quad (4)$$

The summations are carried out over the range of possible velocities. Of course the displacement vector found by these expressions are not exact. If the errors are independent and additive, a covariance matrix S_c can be associated to the estimates and used as a measure of the confidence in the estimate.

$$Sc = \begin{bmatrix} E\{(u-u_c)^2\} & E\{(u-u_c)(v-v_c)\} \\ E\{(u-u_c)(v-v_c)\} & E\{(v-v_c)^2\} \end{bmatrix} \quad (5)$$

with

$$E\{x\} = \frac{\sum_u \sum_v R_c(u,v)x}{\sum_u \sum_v R_c(u,v)} \quad (6)$$

A.2 Neighbourhood constraint

Suppose that the velocity of each pixel is known in a small neighbourhood around the pixel under consideration. If all these velocities are plotted in u,v space, the central pixel is expected to have a velocity "similar" to that of its neighbours. In statistical terms, the velocity of each neighbour can be thought of as being a measurement of the velocity of the central pixel. Of course these measures are not equally important, they should be weighted according to the distance from the central pixel, the larger the distance, the smaller the weight. Specifically a Gaussian mask was used. The weight used for the neighbourhood constrint is called $R_n(u,v)$. The neighbourhood constraint can then be treated in a similar fashion as the conservation constraint, giving $(\overline{u}, \overline{v})$ as the estimated flow field and S_n as the corresponding covariance matrix.

A.3 Combination of both constraints

Let us now represent the velocities as 2×1 vectors. The true velocity at a pixel being U, the neighbourhood estimation $\overline{U}$ and the conservation estimation U_c. From general estimation theory follows that the estimation error is given by:

$$\iint \begin{array}{l} \left[(U-\overline{U})^T S_n^{-1}(U-\overline{U}) \right. \\ \left. +(U-U_c)^T S_c^{-1}(U-U_c)\right] dxdy \end{array} \quad (7)$$

An iterative scheme is used to minimise this error.

$$\begin{array}{ll} U^0 & = U_c \\ U^{k+1} & = [S_c^{-1} + S_n^{-1}]^{-1}(S_c^{-1}U_c + S_n^{-1}\overline{U}^k) \end{array} \quad (8)$$

A.4 The multi-scale approach

If a multi-resolution pyramid is constructed from the images, starting the optic flow calculations at the coarsest level reduces the search space. This is especially important if temporal sampling was only very sparse. Therefore two types of multi-resolution pyramids were examined, the Gaussian pyramid and the Laplacian pyramid ([3]). In [1] the grey levels in the Laplacian pyramid are used as conservation constraint. We obtained better results using a combination of grey level, gradient and texture measures in a Gaussian pyramid.

B. *Warping of the images*

Once the flow field is accurately estimated between each pair of time-sequential images, the images can be warped. This is done by bilinear interpolation.

III. The adaptive temporal low-pass filter

If a simple temporal low-pass filter is used to reduce the noise, non-stationarities in the image sequences that can not be accounted for by motion will still be blurred. Therefore, a locally adaptive version of the temporal low-pass filter ([4]) is used. If d is the number of images to be used in the noise reduction and $C_i(x,y)$ the spatial correlation between a small rectangle centered at point (x,y) in image k and the corresponding rectangle in image $k-i$, the grey value $\hat{G}_k(x,y)$ at that position in the filtered image number k is given by:

$$\begin{array}{ll} \hat{G}_k(x,y) & = \left[1 - \sum_{i=1}^d \alpha_i(x,y)\right] G_k(x,y) \\ & + \sum_{i=1}^d \alpha_i(x,y) G_{k-i}(x',y') \end{array} \quad (9)$$

$$\alpha_i(x,y) = \begin{cases} \frac{1+C_i(x,y)}{4d} & if\ C_i(x,y) > Threshold \\ 0 & otherwise \end{cases} \quad (10)$$

$\hat{G}_k(x,y)$ is the estimation of the k^{th} image of the sequence, (x',y') are the coordinates in image $k-i$ corresponding to the coordinates (x,y) in image k. The size of the correlation window and the value of the threshold are the two adjustable parameters of the filter.

IV. Results and discussion

A. *Results*

The method was tested on a sequence acquired from a moving car and which was sampled at about 2 frames per second. Before processing, zero mean additive Gaussian noise with a standard deviation of 20 was added. Figure 1 shows two time-sequential images of the sequence.

To evaluate the results, the signal-to-noise ratio was estimated before and after restoration. This was done at several values of the threshold and the correlation window size. The SNR was estimated as:

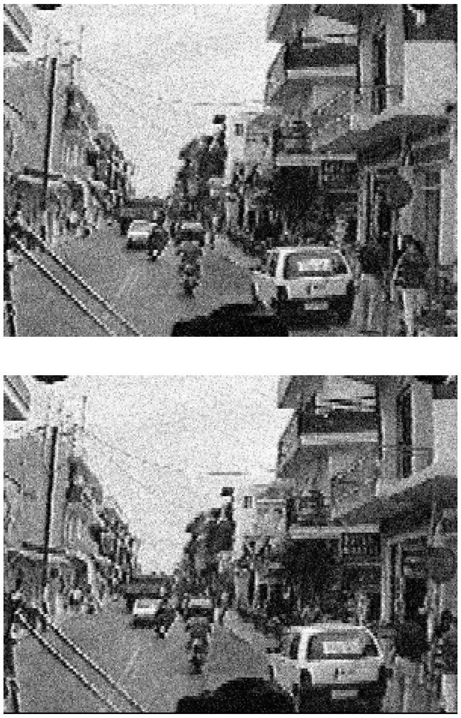

Fig. 1. Example of two time sequential images

MET	TH	CW	SC	RMSE	SNR
ORIG	0	0	1	0	35.9
DEGR	0	0	.96	18.9	12.9
TA	0	0	.784	43.7	24.9
MC	0	0	.945	22.8	25.6
AF	-1	5	.971	16.7	21.4
AF	-1	10	.971	16.6	21.2
AF	-1	20	.972	16.6	21.1
AF	-1	30	.972	16.6	21.1
AF	0	5	.971	16.8	21.4
AF	0	10	.971	16.8	21.2
AF	0	20	.971	16.7	21.1
AF	0	30	.971	16.6	21.1
AF	.5	5	.970	17.1	20.6
AF	.5	10	.970	17.1	20.8
AF	.5	20	.970	17.1	20.7
AF	.5	30	.970	17.1	20.8
AF	.9	5	.967	17.8	19.1
AF	.9	10	.967	18.2	16.4
AF	.9	20	.967	18.2	15.0
AF	.9	30	.967	18.2	14.9

TABLE I

Overview of the quality of the restoration

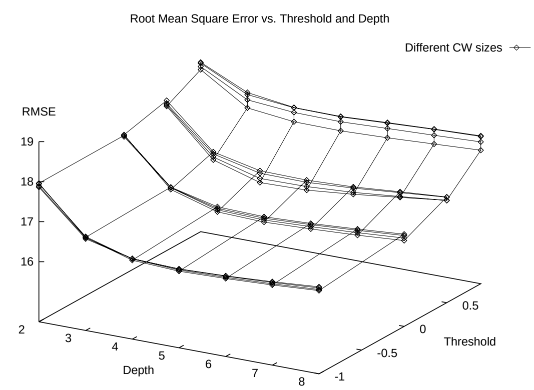

Fig. 2. Influence of Threshold and Depth on RMSE

$$SNR = 10 \times log_{10} \left[\frac{\sigma^2_{max}}{\sigma^2_{min}} \right] \qquad (11)$$

σ^2_{max} and σ^2_{min} are respectively the largest and smallest variance of the grey level distribution calculated on a 10 by 10 square scanning the whole image. The error measures that were used to assess the quality of the restoration are the absolute mean error (AME), the root mean square error (RMSE) and the maximum absolute error (MAE). Also, the spatial correlation (SC) between both images was included in the comparison. Table I shows the results for different values of the system parameters. Because the three error measurements (AME, RMSE and MAE) show a similar behaviour, only RMSE is included in the table. The column labeled MET gives the method used in the noise reduction. The first two rows in the table present measurements on the original image and the degraded image respectively. "TA" means that temporal average is applied on the raw images (without any motion compensation). "MC" is temporal average after motion compensation and "AF" is the result of the adaptive filter. The depth , i.e. the number of subsequent images considered in the restoration was eight in this experiment. The columns labeled "TH" and "CW" respectively show the threshold and correlation window size used for the adaptive filter.

Because of the threshold, only some pixels will be changed during the restoration. In table II the percentage of changed pixels for the test sequence is shown in function of depth in the image sequence. Results are given for different values of the correlation window size (CW) and the threshold (TH).

Figure 2 shows the Root Mean Square Error as a function of both threshold and image sequence depth. In this plot values of different sizes of the correlation window are shown simultaneously.

B. Discusion

From table I can be inferred that a temporal averaging over some images of the sequence without any motion compensation doubles the signal-to-noise ratio. However, error measurements between the original and the "restored" image are very poor. This shows that in evaluating a noise reduction technique used for image restoration, measurements of SNR and error rates need both to be taken into account. For the temporal averaging after motion compensation but without the

TH	Depth	CW = 5	CW = 10	CW = 20
−1	2	91.2	91.2	91.2
	3	80.3	80.3	80.3
	4	69.4	69.5	69.4
	5	64.2	64.3	64.3
	6	56.3	56.3	56.4
	7	50.9	51.0	51.0
	8	48.2	48.3	48.3
0	2	85.1	88.8	90.1
	3	70.8	75.4	78.4
	4	56.8	59.8	62.2
	5	49.4	52.4	55.0
	6	42.3	44.5	47.3
	7	37.4	38.2	40.6
	8	35.3	36.9	39.0
.5	2	72.5	78.2	83.3
	3	58.2	62.3	69.7
	4	43.2	42.5	43.2
	5	37.0	34.3	32.4
	6	32.1	29.9	27.7
	7	26.0	24.2	22.1
	8	25.3	24.1	21.0
.9	2	43.6	38.4	34.4
	3	33.1	28.8	25.5
	4	22.8	17.3	14.1
	5	20.9	15.9	10.9
	6	18.2	13.8	9.9
	7	14.5	10.0	6.0
	8	13.4	8.7	5.8

TABLE II

Percentage of points considered in the restoration as a function of the depth in the image sequence for different values of threshold (TH) and correlation window size (CW)

adaptive filter, the SNR is still doubled and spatial correlation is comparable to the values of the degraded image. However, the RMSE is still larger between the restored image and the original than between the degraded image and the original. Using the adaptive temporal average improves the SNR (by a factor 1.6) and spatial correlation while reducing the error rate.

Figure 2 shows that using more than four subsequent images in the adaptive filter does not reduce RMSE any further. This is probably due to the low correlation found in the adaptive filter as images get further apart in the sequence.

V. Conclusion

A multi-resolution optic flow estimation method was developed to compensate for motion in image sequences with sparse temporal sampling. After motion compensation an adaptive temporal lowpass filter was used to reduce the noise in the sequence while preserving the non-stationary parts of the images. The algorithm was tested on an image sequence acquired from a vehicle moving through a small town. The images were degraded by zero-mean Gaussian noise. Results of the adaptive temporal lowpass filter show an increase of the signal-to-noise ratio while improving spatial correlation and error rates between the restored image and the original.

References

[1] A. Singh, *Optic Flow: A Unified Perspective.* Washington: IEEE Computer Society Press, 1991.

[2] R. Haralick, “Statistical and structural approaches to texture,” *Proc. IEEE*, vol. 65, no. 5, pp. 786–804, 1979.

[3] P. Burt, “Multiresolution techniques for image representation, analysis and 'smart' transmission,” *SPIE Visual Communications and Image Processing*, vol. 1199, pp. 2–15, 1989.

[4] R. Kleihorst, R. Langendijk, and J. Biemond, “Noise reduction of image sequences using motion compensation and signal decomposition,” *IEEE-IP*, vol. 4, no. 3, pp. 274–284, 1995.