Using Optical Flow to Reduce Noise in Image Sequences

Roman Dudek, Francisca Quintana, and Carmelo Cuenca
Departamento de Informática y Sistemas, Universidad de Las Palmas de GC., Las Palmas de GC., Spain

Abstract— *In this paper we propose a method for noise reduction in image sequences, based on the optical flow, consisting in tracking each pixel's position in the previous and following images. In some cases the optical flow imperfections can cause artifacts. To prevent them we also propose an improvement to the method based on the estimation of the imperfections of the optical flow. Using that estimation we adaptatively choose either a temporal or a spatial based noise reduction algorithm to be applied in different image zones. Our results show an important noise reduction, even with complex image sequences.*

Keywords: Optical Flow, Noise Reduction

1. Noise in Image Sequences

Noise is an intrinsic characteristic of image sequences [1]. Either we use digital or chemical resources for shooting image sequences, an important amount of noise is introduced due to the physical nature of the shooting process [2]. Generally, the level of noise produced during this process is fairly constant and keeps at a fairly low level. In normal situations, when filming in good light conditions, the level of the signal vastly outweights the level of noise. In other words, the signal to noise ratio is very high, and consequently noise is not a problem. However, when the acquisition conditions worsen, for example because of the falling of the light levels, the levels of the signal drops nearer to the constant noise level, thus producing a lower signal to noise ratio. This noise needs to be frequently reduced to make further digital processing of the film image sequences. Although many people consider that they miss something important if noise is removed from the grainy film look of the movie theater projections, noise is a technical problem that must be reduced because of their interferences with the digital postproduction process. The noise can be easily added after the image processing, if desired so, for artistic reasons.

2. Usual Methods of Noise Reduction and Our Proposed Method

Multiple methods have been proposed and used for reducing noise. The most obvious approach to reduce the image noise is by suppressing the high frequencies in the image spectrum characteristic, for example with Gaussian filtering [3]. While this technique reduces the image noise, also some important details of the image would be removed. Adaptive methods have also been developed, aimed at distinguishing for example the image edges and do not blur across them [4] [5]. However, every such method has its limitations: For example, the edge detecting method would blur any detailed textures with not enough contrast to be classified as edges. Another approach is based on averaging multiple adjacent images of an image sequence. However, this method only works well for static scenes as any moving object in the scene creates ghost-like trails. Other works, also based on using temporal filtering, have addressed the problem of moving objects in image sequences [6] [7].

In our work, we propose and test a noise reduction method based on averaging the image pixel values along the sequence time. As in the simple frame averaging method described above, we will use the sequence time consistency to distinguish noise from actual image. However, to extend the method's applicability to moving objects, we will use the optical flow (OF) to track each pixel's actual position in previous and following images. In the image zones where the OF fails we apply spatial filtering instead of temporal filtering.

3. Optical Flow Based Noise Reduction

In our process, we first calculate the OF fields[9][10][11] for each pair of consecutive images of the sequence, for both forward and backward directions of the OF. Once the OF fields are calculated, we perform the actual image averaging: for each pixel of a given image, we track the corresponding pixel coordinates in certain number of images preceding and following the current image in the sequence[12][13]. The coordinates are calculated recursively, following the track of the pixel along the sequence. Then, the resulting pixel value is calculated as the average value of the tracked pixels:

$$I_{average}(x,y) = \frac{\sum_{n=-r}^{r} I(x_n, y_n)}{2r+1} \qquad (1)$$

Selecting the value of r (number of frames we use for averaging) requires a compromise as increasing its value increases the grade of noise reduction but it also increases the demands on the OF precision, as inaccuracies in the fields would accumulate while tracking individual pixels along more frames. Our tests show that the best results are obtained by averaging 3 to 7 frames, i.e., $r \in [1,3]$.

The main limitations of the method are the precision of the OF fields and the occlusions or large changes in illumination in the image sequence, which can prevent us from following a given pixel position over long sequences of images.

As a first measure to minimize these problems, during the calculation of the noise reduced current frame, we use a group of images consisting of both the previous and the following frames, with the current frame in the center of the group. This leads to a lower error accumulation than, for example, only using frames which are previous to the current one.

Even with perfect OF fields available, the time averaging method will fail in zones of occlusions. Practically, the OF obtained by an estimation method will contain certain errors, with problematic zones containing large errors. Experiments show that ignoring such errors can cause important artifacts in zones of the scene where the OF field is incorrect (or even undefined in case of occlusions or transparencies). We developed a method to detect such zones and threat them differently.

3.1. Detecting and Treating Zones of Large Errors in OF

A good detection of the OF validity is needed in order to distinguish where OF based method is applicable and where intra-frame filtering should be used instead.

We can assume that where the OF vectors got "lost", and could not correctly follow the movements in the scene, the values of pixels that we find in the neighboring frames using these OF vectors will differ. To estimate these OF errors for each result pixel, we propose to calculate the dispersion (medium square error) of the values used for the calculation of the average value of each result pixel:

$$E(x,y) = \frac{\sum_{n=-r}^{+r} \left(I_n(x_n,y_n) - I_{average}(x,y)\right)^2}{2r+1} \tag{2}$$

Unfortunately, considering that the number of averaged values in the above sums is relatively low (3 to 7), the error measure $E(x,y)$ itself contains a large amount of noise.

To analyze the properties of $E(x,y)$, we can express it as a sum of the "OF Error" $E_{OF}(x,y)$, caused only by the OF imperfections, and a noise component $N_{error}(x,y)$:

$$E(x,y) = E_{OF}(x,y) + N_{error}(x,y) \tag{3}$$

The amplitude of $N_{error}(x,y)$ can be derived from the amplitude of the noise in the source images. Let the individual source image pixels be considered the sum of the "Signal" and a "Noise":

$$I_n(x,y) = S_n(x,y) + N_n(x,y) \tag{4}$$

Let us suppose that the OF vectors were perfect, so the $S_n(x,y)$ are equal for each n from $[-r,+r]$, and $E_{OF}(x,y) = 0$. Then,

$$\begin{aligned} N_{error}(x,y) &= E(x,y) \\ &= \frac{\sum_{n=-r}^{r} \left(I_n(x_n,y_n) - I_{average}(x,y)\right)^2}{2r+1} \end{aligned} \tag{5}$$

The value $N_{error}(x,y)$ would converge to a certain constant c for large values of r, with a high number of averaged frames. This constant could be used as a threshold to decide the validity of the OF fields for a certain pixel: If the error $E(x,y)$ is similar to c, the pixel was likely correctly followed. If the $E_{OF}(x,y)$ is much higher than c, the OF is likely invalid for this pixel, and time averaging should not be used. We use simple spatial Gaussian filtering of the image for the given pixel instead:

$$I_{result} = \begin{cases} (g \circ I_0)(x,y) & \text{for} E(x,y) > c \\ I_{average}(x,y) & \text{for} E(x,y) \leq c \end{cases} \tag{6}$$

However, as we need to use only low values of r ($r \in [1,3]$), the $E(x,y)$ values do not converge enough to the actual noise level of the sequence. There is an important noise component in our $E(x,y)$ itself, making impractical such direct decision per pixel. In practice, the random dispersion of the total error measure $E(x,y)$ is frequently larger than the $E_{OF}(x,y)$ we are actually trying to detect, so neighboring pixels would be frequently randomly misclassified due to this noise component.

To make a better classification, we need to reduce the randomness in the $E(x,y)$ error measure field. We propose to carry out a spatial averaging of the $E(x,y)$ field. This spatial averaging, added to the temporal averaging used to create the $E(x,y)$ itself, can largely reduce the randomness of the $E(x,y)$ error measure, while not affecting much the $E_{OF}(x,y)$ component that we are trying to detect, supposing that it is locally smooth anyway. We used Gaussian filter for this averaging operation, to create a filtered error measure $E'(x,y)$

$$E'(x,y) = (g \circ E)(x,y) \tag{7}$$

The modified algorithm to obtain the final result will use $E'(x,y)$ instead of $E(x,y)$:

$$I_{result} = \begin{cases} (g \circ I_0)(x,y) & \text{for} E'(x,y) > c \\ I_{average}(x,y) & \text{for} E'(x,y) \leq c \end{cases} \tag{8}$$

Using $E'(x,y)$, our threshold classification method provides a much better detection of the problematic zones. However, the application of a thereshold classification causes some visible artifacts along the edges of the zones where the decision changes. In order to reduce such artifacts, we propose to create a thin transition zone using a clamped blending equation instead of a thresholding:

Fig. 1: Sample image from the original sequence.

Fig. 2: Sample image after simple frame averaging.

$$p = clamp((E'(x, y) - c) * s) \quad (9)$$

where s is a user defined constant and $clamp()$ is a function limiting p to the range $[0, 1]$. Then, the final result is obtained by using a blending equation instead of using a threshold:

$$I_{result}(x, y) = p * (g \circ I_0)(x, y) + (1 - p) * I_{average}(x, y) \quad (10)$$

In our test application, the constant s is a user defined value adjusted in such a way that a transition zone of only a few pixels wide will be created around OF error zones. Similarly, c is adjusted by the user in order to correctly detect the zones of OF errors, while ignoring errors too small to cause visible artifacts in the result. The adjustment of other parameters, like the Gaussian filtering radius, depend on the resolution and noise level of the used image sequence. However, once these values are adjusted, they seem to be constant for all shots from the same original negative film roll, so only few adjustments need to be done for each film project.

4. Results

For our tests we used a variety of digitalized image sequences, originally shot on negative film. We have used that kind of sequences as they present an important ammount of noise, so they can take advantage of the proposed method. We obtained important noise reductions without observably suppressing any detail in the scene. Figure 1 shows a sample image from one of the original image sequences.

In Figure 2 we can see the result obtained after a simple frame averaging of five frames. We can see that the image detail is lost due to the motion trails created around any object in movement. These trails can be clearly observed in the heads of both men, and in the left shoulder and the hand of the postman, which are the parts with large movements along the five frames. The noise level is generally reduced, as can be obviously observed in the background, for example. However image stays sharp only in the few parts where there was no movement in the range of the five frames used.

After applying a simple Gaussian filtering to the image we obtain the image shown in Figure 3. The Gaussian filter radius was set just large enough to provide a visually similar noise reduction as it would be obtained by averaging five frames of the sequence. We can see that there are no trails around the shoulder, as in Figure 2, but edges are visibly blurred. Compare the trousers belt edges and the shirt wrinkles of the postman, for example. Some image detail is visibly lost.

Figure 4 shows the frame averaging of 5 frames using OF based tracking of the pixel positions. We can appreciate that noise is reduced and there are no trails of simple frame averaging. There is no detail loss of the spatial Gaussian filtering. However, artifacts can be found in some zones. For example, looking at the hand of the postman, the details of the fingers are completely lost. Moreover, a black strip separating the shirt and the arm has appeared, thus looking like the arm is not overlapping the shirt but behind it.

Fig. 3: Sample image after applying a simple Gaussian filter.

Fig. 4: Sample image after averaging using OF based, with pixel following over 5 frames.

Fig. 5: Sample image using our adaptive algorithm. A simple, hard threshold decision is used in this case.

Fig. 6: Mask image resulting from a simple threshold applied to detect errors in OF.

Fig. 7: Sample image using our adaptive algorithm. Improved threshold is used.

Fig. 8: Mask image resulting from the improved threshold, using a filtered error measure and smooth transitions.

Figure 5 is the result of our adaptive algorithm based on the detection of large errors in the OF. The OF corrected frame averaging is used where errors are small. Gaussian filtering is used where OF errors are large. Although no artifacts can be observed here, however the hard threshold decision based on unfiltered error measure can make spurious spots appear in some scenes, which can be clearly appreciated in Figure 6. This figure shows the mask resulting from a simple threshold applied to detect errors in OF. White pixels represent the zones of large error, where Gaussian filtering image will be used. Black pixels represent the zones of apparently correct OF, where the preferable OF corrected time averaging will be used. We can observe that there are many white isolated points which make random noise appear in an image sequence.

In order to avoid these artifacts we applied the improved threshold of our proposal, with filtered error measure and smooth transition between the zones, whose result can be seen in Figure 7. In that figure it is very difficult to find any visible statical artifact in the image, as can be seen in the mask represented in Figure 8. We can observe in the mask that the spurious spots which caused random noise have been eliminated due to the error measure filtering. Moreover, softening the outlines, using a soft threshold instead of a hard one, further reduces the visibility of any spot artifacts left. As a result, the resulting image preserves, even makes clearer, any detail present in the original sequence, while reducing the noise.

5. Summary

In this paper we have proposed a method for noise reduction in image sequences based on the optical flow. It consists in calculating the OF fields for each pair of consecutive images of the sequence, for both forward and backward directions of the OF. Then we perform the actual image averaging. For each pixel of a given image, we track the corresponding pixel coordinates in certain number of images preceding and following the current image in the sequence. The coordinates are calculated recursively, following the track of the pixel along the sequence.

The main limitations of the method are the precision of the OF fields and the occlusions or large changes in illumination in the image sequence, which can prevent us from following a given pixel position over long sequences of images. Experiments show that ignoring such situations can cause important artifacts in zones of the scene where the OF field is incorrect (or even undefined in case of occlusions or transparencies). To prevent these artifacts we also proposed an improvement to the method based on the estimation of the imperfections of the optical flow. Using that estimation we propose to adaptatively choose either a temporal or a spatial based noise reduction algorithm to be applied in different image zones, with good transitioning on the borders between the zones. Our results show an important noise reduction, even with complex image sequences.

References

[1] Russ J.C. The Image Processing Handbook, CRC Press, 2006.

[2] Reibel Y., et al. CCD or CMOS Camera Noise Characterisation. In Proceedings of the European Physical Journal of Applied Physics, 75–80, 2003.
[3] Gonzalez R.C. and Woods RE. Digital Image Processing. Prentice-Hall, Englewood Cliffs, NJ. 2002.
[4] Perona P. and Malik J. Scale Space and Edge Detection Using Anisotropic Difussion. IEEE PAMI vol.12, No.7 629–639, July 1990.
[5] Lee S. and Kang M. Spatio-Temporal Video Filtering Algorithm Based on 3-D Anisotropic Diffusion Equation. International Conference on Image Processing. Chicago, IL, USA. 447–450 vol.2, 1998.
[6] Bennett E. and McMillan L. Video Enhancement Using Per-Pixel Virtual Exposures, ACM SIGGRAPH, Los Angeles, California, USA, 845–852, 2005.
[7] Dubois E. and Sabri, S. Noise Reduction in Image Sequences Using Motion-Compensated Temporal Filtering, IEEE Transactions on Communications, 32, 7, 826–831, 1984.
[8] Minelly S., Curley A., Giaccone P., and Jones G.A. Reducing Chromatic Grain Noise in Film Sequences, 1998.
[9] Álvarez L., Weickert J., and Sánchez J. Reliable Estimation of Dense Optical Flow Fields with Large Displacements. International Journal of Computer Vision,39(1), 41–56, 2000.
[10] Beauchemin S.S. and Barron J.L. The Computation of Optical Flow. ACM Computing Surveys, volume 27(3), 433–467, 1995.
[11] Horn B., and Schunk B. Determining Optical Flow. AI Memo 572, Massachusetts Institute of Technology, 1980.
[12] Zhongding Jiang, Tien-Tsin Wong and Hujun Bao. Practical Super-Resolution from Dynamic Video Sequences. In Proceedings of IEEE Computer Vision and Pattern Recognition 2003 (CVPR 2003), Madison, Wisconsin, USA, June 16-22, 2003.
[13] Sean Borman and Robert L. Stevenson. Spatial Resolution Enhancement of Low-Resolution Image Sequences: A Comprehensive Review with Directions for Future Research, Department of Electrical Engineering, University of Notre Dame, 1998.