

Adaptive Spatial and Temporal Prefiltering for Video Compression^{*}

Astrid Lundmark and Leif Haglund

Saab Bofors Dynamics AB
firstname.lastname@dynamics.saab.se

Abstract. When compressing video sequences, noise is fatal to compression performance. Imaging is inherently a noisy process since the number of photons hitting a detector is a statistical process. When scene illumination is high, the statistical nature of the number of photons hitting each detector is of no importance, since other effects such as quantization dominate. For other applications, military for example, where the imaging has to be done with whatever illumination is naturally available, the noise can be both annoying for the observer and make transmission of the imagery consume excessive amounts of bandwidth. We show results of using content based spatial prefiltering combined with motion vector certainty controlled temporal prefiltering to reduce the noise and thus improve both visual impression and transmission properties.

1 Introduction

There are different sources of noise in digital imagery, including thermal noise caused by electronics, quantization noise, noise caused by imperfections in the detector elements and noise caused by the stochastic nature of the number of photons hitting a detector. The last type of noise is the dominating one for low light levels. For higher light levels this type has less impact since if the illuminance level is multiplied by a factor A , the time mean of the number of photons hitting the detector increases with a factor A , but the standard deviation only increases with a factor of \sqrt{A} . For low light level applications the noise level can be decreased in the registration process by integrating over a larger area, which would decrease spatial resolution, or integrating over a longer time, which would introduce motion blur. A more attractive solution is to perform filtering (in video coding often referred to as prefiltering) to reduce the noise while preserving the content of the scene. Prefiltering is used in many codecs, but the amount of published articles is not proportional to the importance of the subject. A study [1] comparing the effect of different types of prefiltering in combination with MPEG-2-encoding investigates Wiener filtering, wavelet based filtering, an ordered statistics filter and using the MPEG-2-encoder as prefilter by setting low transform coefficients to zero. The wavelet filter showed best results of the filters investigated while the Wiener filter showed the worst results. Another

^{*} This work was partly funded by the Swedish Defence Material Administration.

study [2] uses anisotropic diffusion, both in 2 dimensions and 3 dimensions, as preprocessing for MPEG-4 video encoding. In [3] a rate-distortion optimal choice of filtering and quantization produces a 1 dB improvement in PSNR and significantly improved subjective image quality when encoding a busy but not very noisy sequence at 3.5 Mbit/s using MPEG-2.

The most general form of prefiltering is spatio-temporal prefiltering. To facilitate implementation, the filtering can be divided into spatial and temporal filtering. This approach was taken in [4], and ideas were exploited and extended in [5].

Denoising of still images can be done using the previously mentioned Wiener filtering, ordered statistics filters (median filtering is an example), anisotropic diffusion, or wavelet filtering both in the decimated domain normally used for image coding and in the undecimated (shift invariant) domain [6]. Vector quantization as a means of denoising has been investigated in [7] and also for simultaneous compression and denoising, see [8].

Temporal prefiltering is closely related to motion estimation and compensation. It has sometimes been argued that moving areas need not be conveyed with full spatial resolution since the human visual system cannot detect high spatio-temporal frequencies. In normal viewing situations, however, the fixation point of the eyes can smoothly follow moving objects, so that they will appear stationary on the retina. In most scenarios we cannot monitor the eye movements of the observer, and therefore we have to transmit all areas, moving or stationary, without reducing the spatial resolution. For motion that is hard to track the spatial resolution can be slightly reduced, an approach pursued in [9]. In [10] the motion vectors computed in the hybrid coding loop were used for prefiltering. Up to now the size of the macro blocks sharing a motion vector has been 16×16 , which means that the resolution of the motion vector field is rather low for use in prefiltering. With the coming of the new video coding standard MPEG-4 Version 10, there will be the opportunity for the encoder to adaptively subdivide the block size for motion compensation resulting in less motion fragmented blocks, which should be beneficial for the approach used in [10].

Spatial content based prefiltering works well for removing noise in flat image regions. Temporal motion vector certainty controlled prefiltering works well for removing noise in regions of the image which contain spatial structure. These two types of filtering work by themselves to obtain an image sequence that is both more pleasing to the eye and demands less bandwidth for transmission. Since they have their strengths in different areas of the image, combining the two types of filtering further improves performance.

The content of this paper follows the outline below. The spatial adaptive prefiltering is described in Sect. 2. Temporal adaptive prefiltering is discussed in Sect. 3. Section 4 is devoted to experimental results. An experiment showing the severity of the effect of noise in the input material with regard to compression efficiency is shown in Sect. 4.1. Results of applying the prefiltering is shown in Sect. 4.2. Conclusions are given in Sect. 5 together with suggestions for future work.

2 Spatial Adaptive Prefiltering

The main idea for the adaptive filtering is that visually important regions are characterized as spatially consistent high frequency areas, e.g. edges and/or lines. These ideas have been used for orientation estimation and for image enhancement estimation [11]. In the context of prefiltering for image compression, an important factor to be considered is the computational complexity. The method that will be described below for reducing noise levels in smooth regions while preserving the visually important features, is designed bearing the computational feasibility aspects in mind. The algorithm consists of the following steps:

- Design a filter set consisting of a lowpass filter, LP, and a corresponding highpass filter, HP, in such a way that the sum of them gives an allpass filter.
- Estimate the visually relevant information by
 - Applying a gradient operator estimating δx and δy .
 - Calculating the local energy as $E_{loc} = \delta x^2 + \delta y^2$.
 - Averaging E_{loc} , e.g. by convolution with a Gaussian filter kernel, to get an estimate of the spatially consistent information.
 - Applying a sigmoidal-like function to obtain a weight image.
- Using the information to determine how much of the HP information that should be added to the LP information to produce the output image.

Images from the different steps of the algorithm are shown in Fig. 1 to 3.



Fig. 1. Spatial adaptive filtering. LP image and HP image.



Fig. 2. Spatial adaptive filtering. Weight image used for controlling how much of the HP image that will be added.



Fig. 3. Spatial adaptive filtering. Left: Output image. Right: Input image (“akiyo” with Gaussian white noise, stddev 4, added), for comparison.

3 Motion Vector Certainty Controlled Temporal Prefiltering

Temporal prefiltering has to be motion-compensated in order to prevent smearing of moving objects. For avoidance of artefacts caused by erroneous motion vectors a motion vector certainty measure, see [14], can be used to control the amount of filtering. The certainty measure is low in areas where new image information appears, and more advanced certainty measures give low values also in areas where the motion vector cannot be uniquely determined.

In the special case of the camera being stationary the computationally heavy task of computing a dense motion field can be eliminated, and the assumption made that the motion vector is zero everywhere. The motion vector certainty measure will then react in areas where this is obviously incorrect, e.g. moving areas and areas consisting of uncovered background, see Fig. 4, where an eye-opening has taken place. The prize to be paid for not computing the motion vector field is that moving areas will contain noise that will increase the bit rate. In some applications this is a reasonable trade-off, in others it is not.

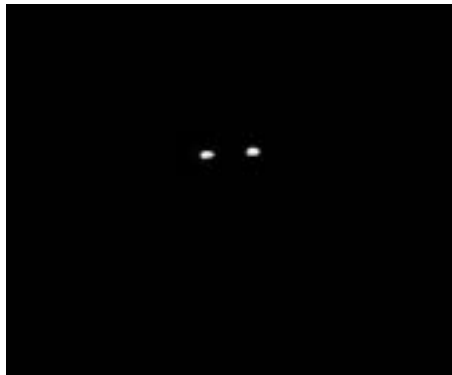


Fig. 4. Motion vector certainty for the “Akiyo” test sequence after an eye opening. Where new image information appears, the certainty measure is low. The motion vector certainty is used to control the amount of temporal filtering.

4 Results

To evaluate the impact of noise on coding performance we conducted an experiment where noise of different strength was added to an input test sequence. These sequences were encoded using a standard codec, and the resulting bitrate was recorded. This experiment is described in Sect. 4.1.

To alleviate the problem of increased bit rate caused by noise in the input sequence, we employed the spatial filtering described in Sect. 2 and the temporal

filtering described in Sect. 3, using the zero vector motion model from Sect. 3, resulting in a computationally lightweight prefiltering system. The results are shown in Sect. 4.2.

4.1 Effect of Input Noise on Compression Efficiency

The bulk of image coding research and development is concerned with input material having relatively high quality, such as the CIF resolution (352×288 pixels) “akiyo” test sequence, which was used in the MPEG-4 development process. To this sequence we added white Gaussian noise with different standard deviation and coded the sequences using MPEG-4-encoding with the first frame coded as an intra (I) frame and the remaining frames using motion compensated prediction (P) frames. The step height of the quantization was set to different values to evaluate coding performance in the range from visually lossless at $Q=4$ to coarse quantization at $Q=28$. The chrominance signals have been set to zero on the input of the codec, using the assumption of a priori knowledge of the signal having zero chrominance, which makes it possible to avoid spending bits on encoding noise in the chrominance bands. The bit rate produced for different combinations of the amount of noise and quantization parameter Q is shown in table 1 and Fig. 5, and the conclusion is that the loss in compression performance due to noise in the input sequence is severe.

Table 1. Number of bits produced for encoding 50 frames (one second) of the “akiyo” test sequence (luminance-only version) with white Gaussian noise having different standard deviation σ added. MPEG-4 was used for coded with the first frame Intra-coded and the remaining frames coded using motion compensated prediction. Larger value of Q means coarser quantization. The data is also visualized in Fig. 5

Q	$\sigma = 0$	$\sigma = 4$	$\sigma = 8$	$\sigma = 16$	$\sigma = 32$
4	463 824	2 756 208	8 324 680	17 566 784	25 839 816
8	193 488	298 736	1 851 304	7 825 256	16 542 568
12	120 416	139 768	315 544	3 745 600	11 654 368
16	93 048	97 992	136 432	1 670 000	8 115 056
20	79 152	82 712	99 952	679 712	5 295 264
24	72 576	75 952	87 408	297 608	3 576 568
28	69 984	73 560	84 808	189 344	2 449 288

4.2 Effects of Prefiltering on Compression Performance

Using no prefiltering, spatial prefiltering, temporal prefiltering, and spatial prefiltering followed by temporal prefiltering, the “akiyo” test sequence with white Gaussian noise with a standard deviation of 4 added was coded with an MPEG-4 codec. Table 2 presents the result for visually lossless quantization produced with

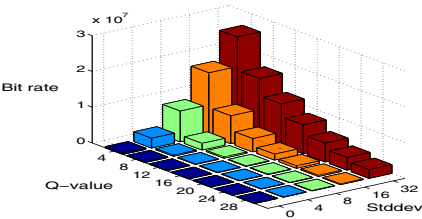


Fig. 5. Number of bits for transmitting one second (50 frames) of the “akiyo” test sequence using MPEG-4 encoding. The Q parameter determines the height of the quantization step, which determines the amount of distortion. The chrominance was set to zero. Table 1 contains the numerical data.

the quantization parameter Q set to 4, and coarser quantization produced with the quantization parameter set to 14. The prefiltering steps do not introduce any objectionable distortion into the sequence. For Q=14, temporal prefiltering by itself is enough to produce a lower bitrate than the bit rate required for coding the “akiyo” sequence without added noise.

Table 2. Using different kinds of prefiltering produces the number of bits shown below for encoding 50 frames (one second) of the “akiyo” test sequence with white Gaussian noise having standard deviation 4 added. The chrominance signals have been set to zero. The Q value determines the quantization step height used in the MPEG-4-encoder.

Q	No prefilter	Spatial prefilter	Temporal prefilter	Spat. & temp. prefilter
4	2 756 208	791 208	1 348 016	454 800
14	114 488	112 568	95 784	95 840

5 Conclusions and Future Work

We have shown that there is a need for means to alleviate the impact of noise for reasons of coding efficiency. Examples of spatial and temporal prefiltering to mitigate the impact of the noise have been demonstrated. For moderate noise levels spatial prefiltering followed by temporal prefiltering performs well, since the spatial prefiltering removes noise from regions where the spatial activity is low and the temporal filtering removes noise from textured regions. When noise levels increase, spatio-temporal filtering integrating motion estimation over several frames might show to be necessary to cope with the increasing difficulty of separating image content and noise.

References

1. van Roosmalen, P. M. B., Kokaram, A. C., Biemond, J.: Noise Reduction of Image Sequences as Preprocessing for MPEG2 Encoding. European Conference on Signal Processing (EUSIPCO) (1998) 2253–2256
2. Tsuji, H., Sakatani, T., Yashima, Y., Kobayashi, N.: A Nonlinear Spatio-Temporal Diffusion and its Application to Prefiltering in MPEG-4 Video Coding. International Conference on Image Processing (2002) 85–88
3. Karunaratne, P. V., Segall, C. A., Katsaggelos, A. K.: A Rate-Distortion Optimal Video Pre-Processing Algorithm. International Conference on Image Processing (2001) 481–484
4. Vasconcelos, N., Dufaux, F.: Pre and Post-Filtering for Low Bit-Rate Video Coding. International Conference on Image Processing (1997) 291–294
5. Gottardo F., Mir, J., Parladori, G., Siracuraza, G. L.: Nonlinear Preprocessing for MPEG-2 Video Sequences. Applied Signal Processing (6):61–70, 1999.
6. Donoho, D. L.: De-Noising by Soft-Thresholding. IEEE Transactions on Information Theory, 41(3):613–627, May 1995.
7. Graña, M., Echave, I. Ruiz-Cabello, J.: VQ Based Bayesian Image Filtering. International Conference on Image Processing (2000)
8. Panchapakesan, K., Bilgin, A., Sheppard, D. G., Marcellin, M. W., Hunt, B. R.: Simultaneous Compression and Denoising of Imagery using Non-Linear Interpolative Vector Quantization. IEEE Digital Signal Processing Workshop (1998)
9. Faber, J., Ihlenburg, L., Meijers, T., Ruschin, D.: Irrelevancy due to Visual Tracking Errors. Picture Coding Symposium (1997)
10. Biloslavo, M., Ramponi, G., Oliveiri, S., Albani, L.: Joint Kalman-Based Noise Filtering and Motion Compensated Video Coding for Low Bit Rate Videoconferencing. International Conference on Image Processing (2000)
11. Bigün, J.: Local Symmetry Features in Image Processing. PhD Thesis, Linköping University, Sweden, 1988. Dissertation No 179, ISBN 91-7870-334-4.
12. Knutsson, H., Wilson, R. and Granlund, G. H.: Anisotropic Non-Stationary Image Estimation and its Applications — part I: Restoration of Noisy Images. IEEE Transactions on Communications, COM-31(3):388–397, March 1983.
13. Haglund, L.: Adaptive Multidimensional Filtering. PhD Thesis, Linköping University, Sweden, 1992. Dissertation No 284, ISBN 91-7870-988-1.
14. Lundmark, A.: Hierarchical Structures and Extended Motion Information for Video Coding. PhD Thesis, Linköping University, Sweden, 2001. Dissertation No.683, ISBN 91-7219-997-0.