A New Cascaded Spatio–Temporal Noise Reduction Scheme for Interlaced Video

Klaus Jostschulte
AG Schaltungen der Informationsverarbeitung
University of Dortmund, Germany
jo@nt.e-technik.uni-dortmund.de

Aishy Amer
INRS-Télécommunications
Montreal, Canada
amer@INRS-Telecom.UQuebec.CA

Abstract

The quality of a television system is significantly determined by the amount of distortions within the displayed image. One of these distortions is white gaussian noise. So a spatio–temporal algorithm for reducing noise in interlaced video is presented in this paper. It consists of a detail preserving spatial algorithm which uses a set of contour oriented lowpass–filters that are controlled via corresponding highpass masks. The influence of the number of different masks as well as of the coefficients are analyzed. The temporal scheme works in two subbands and makes use of some perceiption properties of the human eye. The high efficiency of this cascaded system is proofed by simulations.

1. Introduction

In every visual communication system noise is an unavoidable problem. Especially in television sets noise reduction has to be performed to increase the quality of the displayed image.

In current television sets noise reduction is usually performed by a motion adaptive temporal filter based on a frame or field delay [4]. Because of the interlaced structure of the TV–signal there are problems either in details or moving objects. For this reason a processing of video signals should take care of interlace. This can be achieved taking into account special properties of the human visual system.

A motion adaptive temporal filter has higher efficiency if it is applied to signals that are spatially noise reduced before. But spatial lowpass filters for noise reduction purposes have the tendency to blur edges. In other known algorithms [1, 2] this is avoided by measuring the image activity and switching off noise reduction if the measured value exceeds a threshold that is optimized for a particular signal–to–noise–ratio. The technique that is presented here works without the definition of thresholds and therefore it is appropriate for a wide range of input signal–to–noise–ratios without knowledge of the image statistics.

2. Subband based temporal noise reduction

The task of a noise reduction algorithm is to eliminate a spatially and temporally uncorrelated noise signal from a noisy image to restore the underlying original image. Usually noise reduction is performed by a recursive temporal lowpass filter. To avoid artifacts such as motion blur the filter is controlled by a motion detector which switches off noise reduction in moving image parts. Detailed considerations concerning a motion–adaptive recursive temporal noise reduction are given in [4].

For interlaced video signals a simple recursive filter for noise reduction has a crucial disadvantage. The rasters of consecutive fields do not match (see Figure 1). But a temporal noise reduction has to be performed between pixels of the same position. So there are two possibilities for a temporal filter. The first one is a field based temporal filter. That means the recursive filter is realized with a field delay. In this case the recursive contribution has to be taken from the line above or below or it must be interpolated. The result is a loss of vertical resolution for the output image. The second possibility for the filter is a frame based recursion. The rasters of both images that are used for filtering do match and no loss of resolution is visible. But there is another disadvantage. The temporal difference between the pixels is higher than in case of a field–based filter. Moving objects will produce larger unprocessed areas due to motion detection. The result

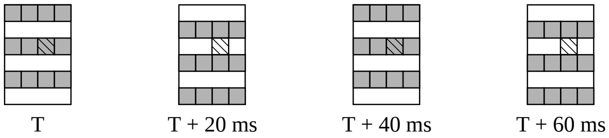

Figure 1: Rasters in consecutive fields do not match in interlaced video

is a reduced efficiency of noise reduction.

The mentioned problems can be solved using a subband based processing adapted to perception properties of the human eye [3, 5, 6]. A general scheme of a subband based processing is depicted in Figure 2.

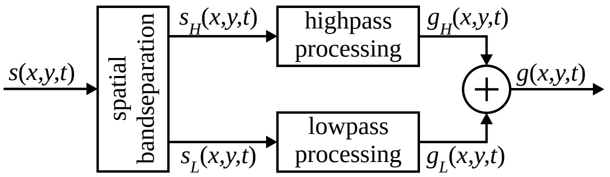

Figure 2: General scheme of a 2–channel processing

According to the perception properties noise reduction in the highpass channel should preferably be frame based to preserve the full spatial reolution. The realized scheme in this channel is given in (1), with k_H the output of the highpass motion detector.

$$\begin{aligned} g_\mathrm{H}(x,y,t) =& k_\mathrm{H} \cdot s_\mathrm{H}(x,y,t) \\ &+ (1-k_\mathrm{H}) \cdot g_\mathrm{H}(x,y,t-T_{frame}) \end{aligned} \tag{1}$$

In the lowpass channel it is desirable to have a firld–based system with respect to motion. But in case of static image content a frame–based method will give best results. Both constraints are fullfilled if a median based deinterlacing is applied. The resulting processing is given in (2), with k_L the output of the motion detector in the lowpass channel.

$$\begin{aligned} g_\mathrm{L} =& k_\mathrm{L} \cdot s_\mathrm{L}(x,y,t) \\ &+ (1-k_\mathrm{L}) \cdot \mathrm{med} \begin{pmatrix} g_\mathrm{L}(x,y,t-T_\mathrm{frame}) \\ g_\mathrm{L}(x,y-1,t-T_\mathrm{field}) \\ g_\mathrm{L}(x,y+1,t-T_\mathrm{field}) \end{pmatrix} \end{aligned} \tag{2}$$

In non–moving homogenous areas the median–filter yields further noise reduction [7]. This has high subjective effect because noise of low spatial frequency is reduced [8].

Results of this temporal filter are significantly better compared to other currently implemented methods. Simulations concerning the signal to noise ratio yields in noise reduction values that are more than 2 dB better than a motion adaptive frame recursion and nearly 3 dB better than a motion adaptive field recursion. Subjective considerations show best noise reduction compared to the other methods without any loss of resolution. A more detailed analyzation of this filter is given in [3, 10].

3. Edge preserving spatial noise filter

In video signals usually spatial correlations exist. In the literature correlation models, e.g. the autoregressive process, are given [9]. By using spatial lowpass filters these correlations are used to preserve the original signal but reduce spatially uncorrelated noise. One simple method for reducing noise is a symmetric 3–tap FIR–lowpass filter with coefficients $\{\frac{1-a}{2}; a; \frac{1-a}{2}\}$. The effect of such a filter on white gaussian noise with input variance σ_in^2 can be determined using (3):

$$\sigma_\mathrm{out}^2 = a^2 \cdot \sigma_\mathrm{in}^2 + 2 \cdot \left(\frac{1-a}{2}\right)^2 \cdot \sigma_\mathrm{in}^2 \tag{3}$$

If the image is unstructured, the original image remains unaffected. So the noise reduction value R can be determined (4).

$$\mathrm{R}\,[\mathrm{dB}] = 10 \cdot \log \frac{\sigma_\mathrm{in}^2}{\sigma_\mathrm{out}^2} = 10 \cdot \log \left(\frac{2}{3a^2 - 2a + 1}\right) \tag{4}$$

In case of edges and lines in the image such a filter will blur them which results in a loss of resolution. In currently implemented methods for spatial noise reduction an image analyzer is used to determine image–activity. Dependent on the output of the analyzer noise reduction is switched off. For such systems particular thresholds are required to determine switching points of the filter. Some disadvantages come up with such a method. First of all not the complete image will be processed because near edges and lines the filter is switched off. Furthermore for the determination of the thresholds a lot of simulations and calculations are required to get optimal results. Besides all thresholds are only valid for a particular signal–to– noise–ratio. For an implementation in a television set this results in further costs for a noise estimation.

Because of the abovementioned problems another concept without thresholds was developed in this work. It is based on the assumption that the image consists of nearly homogenous regions which are separated by edges, lines and corners. Under this assumption noise reduction can be performed in all areas of the image. But the direction of the noise reduction filter must be independently determined for all regions. To avoid blurring, the filter has to be applied along edges and lines but never across them. So for all image structures there exists a lowpass filter which will not filter across structures.

The proposed system works with differently oriented 3–tap highpass filters for analyzing purposes. The filters must have zero response at DC-frequency. This is achieved with filter coefficients of $\{-\frac{1}{4}; \frac{1}{2}; -\frac{1}{4}\}$. The structure orientation is determined by the mask delivering the lowest absolute output.

The lowpass filter for noise reduction is applied with the same mask but using arbitrary lowpass coefficients. The coefficients are chosen to be $\{\frac{1-\mathrm{a}}{2}; \mathrm{a}; \frac{1-\mathrm{a}}{2}\}$. The filter will be analyzed in the following sections, distinguishing between structured and homogenous image regions.

3.1. Noise reduction in structured areas

In structured image areas the direction of the filter is determined by the orientation of image structures. That means that there is only one mask which fits the image structure perfectly. In such cases the chosen mask is uncorelated to noise and no edges will be blurred. So noise reduction in this case can be determined by (4) again.

Figure 3 shows the dependency between the central coefficient a and the noise reduction value R. As it can be seen for a cos^2–shaped filter this yields in a noise reduction of 4.26 dB.

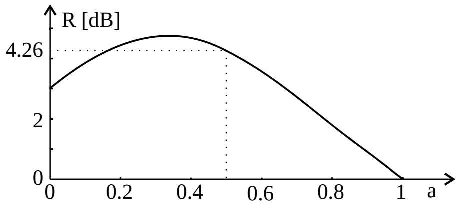

Figure 3: Noise reduction vs. central coefficient

3.2. Noise reduction in homogenous areas

In homogenous regions the behaviour of the filter is determined by noise structures. That means, that the filter will be oriented to structures formed by noise. The result will be a reduced noise reduction value, because mask selection and noise are not uncorrelated.

For the following calculations a cos^2–shaped lowpass filter is assumed.

The noise signal at the input of the filter (signals s_{-1}, s_0, s_1 in Figure 4) is assumed to have a gaussian probability density function (*pdf*) with variance σ^2_{in}.

After multiplication of the central input s_0 with the center coefficient of 0.5 the resulting noise variance σ_0^2 of the intermediate signal w_0 is $\sigma_0^2 = \frac{1}{4} \cdot \sigma^2_{\mathrm{in}}$, so a *pdf* of

$$\mathrm{f}_0(x) = \frac{\sqrt{2}}{\sqrt{\pi} \cdot \sigma_{\mathrm{in}}} \cdot e^{-\frac{2 \cdot x^2}{\sigma^2_{\mathrm{in}}}} \tag{5}$$

results.

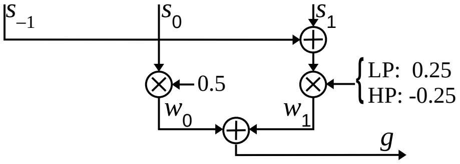

Figure 4: Model of one mask of the analyzing and noise reduction filter

The noise–*pdf* after addition and multiplication of the input signals s_{-1} and s_1 with the the coefficient 0.25 can be calculated to

$$\mathrm{f}_1(x) = \frac{2}{\sqrt{\pi} \cdot \sigma_{\mathrm{in}}} \cdot e^{-\frac{4 \cdot x^2}{\sigma^2_{\mathrm{in}}}}. \tag{6}$$

The different *pdfs* are depicted in Figure 5.

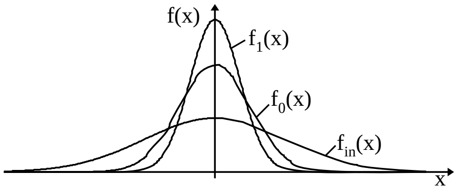

Figure 5: Probability density functions of the intermediate signals

The noise–*pdf* at the output of the filter can be calculated by a convolution of the probability density functions of w_0 and w_1. In case of N masks and with the assumption of independent identically distributed (*iid*) non center inputs the mask with the lowest absolute highpass output is chosen. Regarding to the model of Figure 4, the input mask, which produces the lowest difference between w_1 and w_0, will be the one that is chosen. All other masks will have a higher highpass output and are refused. A typical situation is given in Figure 6. The signal w_0 equals $\frac{x}{2} + x_0$. The signal w_1

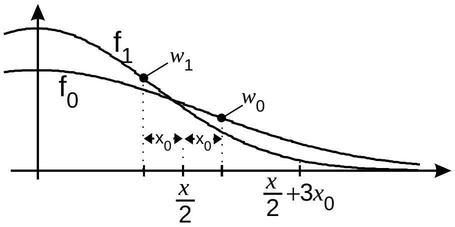

Figure 6: Probability density functions at different steps of the filter

with the lowest difference to w_0 equals $\frac{x}{2} - x_0$. All signals w_1 of the other $N-1$ masks must therefore have samples outside the area $[\frac{x}{2} - x_0; \frac{x}{2} + 3x_0]$.

The probability of a random variable s with variance σ^2 falling into the range $[\frac{x}{2} - x_0; \frac{x}{2} + 3x_0]$ is determined using the gaussian error function erf(·) (7).

$$\begin{aligned} &\mathrm{P}(\frac{x}{2} - x_0 \leq s \leq \frac{x}{2} + 3x_0) = \\ &\quad \frac{1}{2} \cdot \left(\mathrm{erf}\left(\frac{\frac{x}{2} + 3\,x_0}{\sqrt{2} \cdot \sigma} \right) - \mathrm{erf}\left(\frac{\frac{x}{2} - x_0}{\sqrt{2} \cdot \sigma} \right) \right) \end{aligned} \tag{7}$$

With the analyzations above the noise–*pdf* of the filter output g can be determined by the convolution of the probability density function of both branches of the chosen mask multiplied with the probability that the other $N-1$ masks have higher highpass outputs. (8).

$$\begin{aligned} \mathrm{f}_{out}(x) = \int\limits_{-\infty}^{+\infty} &\frac{N \cdot 4}{\sqrt{2\pi\sigma_{\mathrm{in}}^2}} \cdot e^{-\frac{\left(\frac{x}{2}+x_0\right)^2}{\frac{\sigma_{\mathrm{in}}^2}{4}} - \frac{\left(\frac{x}{2}-x_0\right)^2}{\frac{\sigma_{\mathrm{in}}^2}{2}}} \cdot \\ &\cdot \left(1 - \frac{1}{2} \left| \mathrm{erf}\left(\frac{\frac{x}{2} + 3\,x_0}{\frac{\sigma_{\mathrm{in}}}{2}} \right) \right.\right. \\ &\qquad \left.\left. - \,\mathrm{erf}\left(\frac{\frac{x}{2} - x_0}{\frac{\sigma_{\mathrm{in}}}{2}} \right) \right| \right)^{\mathrm{N}-1} dx_0 \end{aligned} \tag{8}$$

With the assumption of a noise output g having zero mean, an output variance $\sigma_{out}^2 = \mathbf{E}[g^2]$ results.

In this way the dependency of noise reduction in homogenous regions and the number of masks can be calculated. The results are given in Table 1 and depicted in Figure 7. It is obvious that the amount of noise reduction decreases with increasing number of masks. So in the end there is a trade–off between structure preserving properties, which requires a large amount of different masks, and noise reduction in homogenous regions.

3.3. Realization of the spatial filter

As mentioned in the last section a compromise regarding the number of masks has to be found. 8 different masks proved to be a good compromise. These masks are depicted in Figure 8. The system uses a 3 × 3–neighbourhood for noise reduction. For this reason it is low expensive concerning memory. Regarding to Table 1 a noise reduction in homogenous areas of 1.51 dB would be expected if all masks have independent non–center inputs. However, because of the special structure of the masks only 4 ones offer independent inputs. A more exact calculation leads to theoretical noise reduction values of 2 dB in homogenous regions.

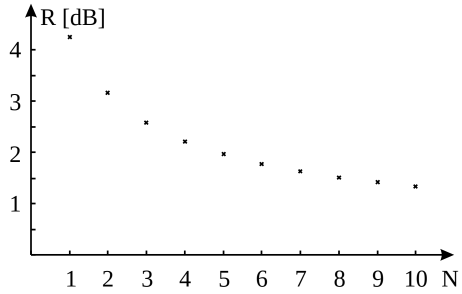

Figure 7: Noise reduction in homogenous regions decreases with the number of different masks

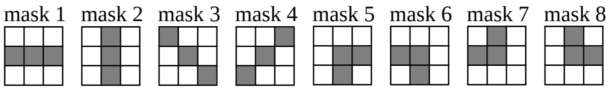

Figure 8: Realized system has 8 masks

Simulations of this filter yields in a gain of the signal–to–noise–ratio of 1 dB up to 2 dB. This proved as sufficient for supporting the described temporal noise reduction. Because of its low expense compared to its efficiency this spatial noise reduction scheme is part of an integrated circuit for receiver based spatio–temporal noise reduction.

4. Results

To proof its performance system simulations of the cascaded spatio–temporal system in comparison to pure spatial and pure temporal noise reduction in presence of white gaussian noise were performed. In Figure 9 the gain in signal–to–noise–ratio of these filters are depicted for four different test sequences.

It can be pointed out that especially at noise levels below 30 dB the results of the complete chain may even be higher than the sum of pure spatial and pure temporal noise reduction. The reason for this is the higher accuracy of the motion detection of the temporal noise reduction.

Furthermore critical sequences for motion adaptive noise reduction (e.g. zoom) deliver good results if they are processed with the combined technique.

Besides the signal–to–noise–ratio does not show the subjective gain that is provided by the visual adaption of the temporal filter.

Table 1: Dependency of noise reduction and number of masks

Number of masks N	1	2	3	4	5	6	7	8	9	10
Noise reduction R [dB]	4.26	3.17	2.58	2.21	1.96	1.77	1.63	1.51	1.42	1.34

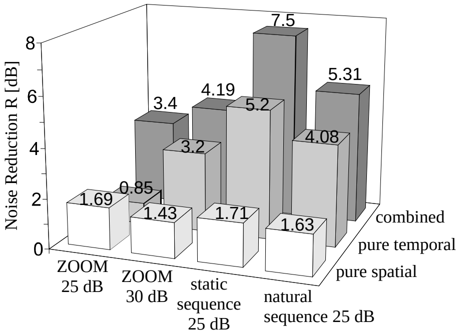

Figure 9: Results of the spatio–temporal system compared to pure spatial and pure temporal noise reduction

5. Conclusions

A cascaded spatio–temporal system for noise reduction in interlaced video signals was presented in this paper. The spatial noise reduction consists of a structure preserving spatial lowpass filter system using complementary highpass filters for image analyzation. The spatial filter works without thresholds and therefore it is appropriate for a wide range of different signal–to–noise–ratios of the input image. It offers good noise reduction with low hardware expense.

The temporal noise reduction works in two subbands and makes use of special properties of the human visual system to perform a temporal noise reduction with high subjective quality.

The cascadation of both systems leads to total results that could be higher than the sum of pure spatial and pure temporal noise reduction. This is always valid if the second system has an analyzation step that is supported by the first system.

6. Acknowledgements

The authors wants to thanks Siemens AG, Munich, Germany, who supported this work within a project concerning receiver–based noise and artifact reduction. Further thanks to Prof. Dr.–Ing. H. Schröder for helpful suggestions and discussions.

7. References

[1] de Haan, G.; Kwaaitaal-Spassova, T.G.; Larragy, M.; Ojo, O.A: *"Memory Integrated Noise Reduction IC for Television"*, IEEE Tr. On Consumer Electronics, Vol. 42, No. 2, 1996

[2] de Haan, G.; Kwaaitaal-Spassova, T.G.; Ojo, O.A: *"Automatic 2-D and 3-D noise filtering for high-quality television receivers"*, International Workshop on HDTV'94,4-B-2,1994-10

[3] Amer, A.; Schröder, H.: *"A New Video Noise Reduction Algorithm* Using Spatial Subbands", Int. Conf. On Electronics, Circuits and Systems; ICECS '96; Rhodes

[4] Drewery, J.O.; Storey, R; Tanton, N.E.: *"Video Noise Reduction"*, BBC Research Department Report, 1984/7

[5] Hiwatashi, K.; Yamaguchi, Y.: *"Perception of Random Monochrome Video Interference"*, NHK Tech. Journal, Vol. 10, No. 5, 1959

[6] Yasuda, M.: *"A Retina Model and its Spatio-Temporal Information Processing"*, NHK Tech. Journal, Vol. 23, No. 5, 1971

[7] Justusson, B.I.: *"Median Filtering: Statistical Properties"*, in Two dimensional digital signal processing II, T.S. Huang (editor), Springer Verlag, 1981, pp. 161-196

[8] Fujio, T.: *"A Universal Weighted Power Function of Television noise and Application to High–Definition TV System Design"*, IEEE Transactions on Broadcasting, Vol. 26, No. 2, 1980

[9] Jain, A.K.: *"Fundamentals of digital image processing"*, Prentice–Hall, 1989

[10] Jostschulte, K.; Amer, A.; Schu, M.; Schröder, H.: *"A Subband Based Spatio-Temporal Noise Reduction Technique for Interlaced Video Signals"*, IEEE International Conference on Consumer Electronics, June 2-4, 1998, Los Angeles