GRAIN NOISE REDUCTION USING THE HUMAN VISUAL SYSTEM CHARACTERISTICS

Valery Naranjo, Soledad Gómez, Antonio Albiol, Jose M. Mossi

Comunication Department, Politechnic University of Valencia
Camino de Vera S/N, 46020, Valencia, Spain
phone: + (34) 963877746, fax: + (34) 963877309 your fax, email: vnaranjo@dcom.upv.es , sogogar@iteam.upv.es

ABSTRACT

This paper deals with grain noise artifact reduction on archived films. Concretely, our research has been focused on finding a technique that let us not only reduce the grain noise but also preserve the original image quality as accurate as possible. The paper investigates the influence of different types of noise reduction filters in order to introduce the problem. Moreover, for each technique some conclusions are exposed. Finally, a linear spatial-temporal technique is proposed and described. This technique adapts a spatial-temporal response according to the human visual system behavior. Taking advantage of it, we get great results reducing the grain noise of several image sequences and preserving their original quality as much as possible. Moreover, the filter does not need motion estimation and it is implemented using separable filters, resulting a computational efficient implementation.

1. GENERAL INFORMATION

50% of the films recorded before 1950 have been lost forever. That is why image digitization and restoration turn out to be such important activities nowadays. All of these films suffer from at least one of the following degradations: flicker and local brightness variations, full damaged frames, scratches, dust and dirt spots, vibration and grain noise. In this paper, we will focus on the grain noise problem, which is inherent to the photographic material where the film was originally recorded. To get rid of this noise, many filtering techniques have been proposed based on the computation of a mean intensity value on a local neighbourhood at each image pixel. This neighbourhood, in the particular case of video sequences, includes both spatial and temporal pixels close to the original one. However, these techniques tend to displace structures and blur their boundaries. These collateral effects should be minimized as much as possible. For these reasons, our main research has been centered not only on finding a technique that let us reduce the grain noise but also preserve the original subjective image quality, reducing as much as possible the blurring effect.

A review of the existing techniques is presented in section 2. In section 3 the proposed filter is presented and described. Some important results are shown in section 4. We conclude this paper in section 5.

2. STATE OF THE ART

The noise sequence model to be used can be defined as:

This work has been supported by the Polytechnic University of Valencia interdisciplinary project 5607-2004 and the Ministerio de Educacin y Ciencia-FEDER program under grant TEC 2005-07751-C02-01.

$$g(x,y,t) = f(x,y,t) + n(x,y,t) \tag{1}$$

where g(x,y,t) and f(x,y,t) represent the pixel intensity levels of the original and processed images, respectively, for the t-th frame in the spatial position (x,y), and n(x,y,t) represents the stationary noise component. The main aim will be to recover f(x,y,t) from g(x,y,t) as accurate as possible.

For that, temporal or spatial-temporal filters are commonly used. Spatial-temporal techniques reduce more noise than temporal filters, taking advantage of spatial and temporal correlation, but suffer from a high computational cost. Both of them, could cause blurred border, in temporal direction, in the presence of motion. Several alternatives (linear and non-linear) to solve this problem have been proposed such us to use filters with adaptive coefficients or motion compensation filters.

A review of these techniques is discussed in next points.

2.1 Average Filters

Several spatial-temporal filters proposed in the literature have been implemented generalizing 2D filter techniques, adding the temporal dimension as follows:

$$\widehat{f}(x,y,t) = \sum_{p,q,t} w(p,q,t)g(x-p,y-q,t-l) \tag{2}$$

The most accurate implementation of this filter is to use coefficients with the same weight, $w(x,y,t) = \frac{1}{MNL}$, where M, N and L are the size for each direction. However, this technique suffers from the blurring problem in those image areas with high motion or spatial borders. There are several techniques proposed in the literature with the aim of solving this problem, which can be classified as: finite response filters (FIR) and infinite response filters (IIR).

2.1.1 FIR Filters

The coefficients w(p,q,t) can be stationary or adaptive and are chosen to get a concrete objective. The most common option is to reduce the quadratic mean error $min_{w(p,q,t)}E[(f(x,y,t)-\widehat{f}(x,y,t))^2]$. In this case, the equation 2 becomes a Wiener Filter. Ökan et al. [1] present an efficient algorithm to implement it, Chan et al. [2] propose a different solution modelling the non-stationarity of the sequence. Compensated motion versions of the above comment filters can be found too, such as the one proposed by Boyce [3]. In order to solve the blurring problem, causes for reasons other than motion, for example spatial filtering, the coefficients are adapted according to spatial correlation. All these techniques provide best results in comparison with the temporal filter without motion compensation or the motion compensated filter without non-adaptive coefficients, but, anyway, the blurring effect continues being so appreciable.

2.1.2 IIR Filters

Both temporal and spatial-temporal IIR filters, have the trade off between the quantity of noise reduction and the memory needed for their implementation. The recursive IIR filters, solve this problem as follows:

$$\widehat{f}(x,y,t) = [1-\alpha(x,y,t)]\widehat{f_b}(x,y,t)+\alpha(x,y,t)g(x,y,t) \quad (3)$$

where $\widehat{f_b}(i,j,t)$ is the prediction of the original sequence before being actualized by the filter, and $\alpha(x,y,t)$ is the filter gain. Depending on the way we use to predict $\widehat{f_b}(x,y,t)$ or actualize the gain, we can found a great deals of different techniques, e.g the proposed by Kleihorst et al. [4].

2.2 Ordered-Statistic Filters

The intensity levels of pixels are ordered inside to the analysis window, before operating with them. The most commonly used are the median filters. The main problem is that narrow objects with fast motion are eliminated. Multilevel or multistage filters are proposed in order to reduce this problem, e.g we can cite Alp and Neuvo [5].

2.3 Bayesian Filters

The technique consists of maximizing a likelihood function, $\widehat{f}(x,y,t) \leftarrow max_{f'(x,y,t)}\, p\{g(x,y,t)|f'(x,y,t)\}$. The compensated motion filter has an important problem: the final result depends on the precision of the motion estimator. If the level of noise is high, the motion estimator could not be very precise. In the literature we can found several Bayesian filters which could combine the noise reduction and motion estimation in an only step, eg. we can cite Brailean and Katsaggelos [6] who propose a pel-recursive motion estimator.

2.4 Multi-resolution Filters

These techniques become the most popular applied to image sequences, in several application such us video compression. The principal idea is to decompose the 2D signal in different bands with different resolution in order to sparse the noise in every band, whereas the signal energy remains centered in a few bands. Different techniques have been proposed, the most popular is the one called Coring, which uses de advantages of the DWT, e.g Roosmalen et al. [7].

3. THE SPATIAL-TEMPORAL LINEAR FILTER

Our main research has been centered not only on finding a filter which let us reduce the grain noise but also preserve the original subjective image quality, reducing as much as possible the blurring effect. The implemented filter consists of a linear FIR filter which uses de human visual system properties in order to introduce degradations, unavoidable in the filtering process, only when the human system is not able to appreciate them. Moreover, it is very important to stress the importance of its separability in order to achieve a high efficient implementation.

3.1 Basic Idea

The spatial-temporal filter proposed in this paper, is based on the spatial-temporal response of the Visual Human System [8] (SVH), see figure 1.

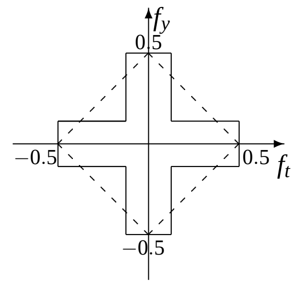


Figure 1: Spatial-temporal response of visual human system (- - -) and implemented filter response (——).

By seeing this response (- - -), we can emphasize the next points:

- The human eye is not able to detect simultaneously high spatial and temporal frequencies. We are not able to see in detail objects whose motion is very fast. On the other hand, objects with slower motion can be better appreciated.
- Our spatial-temporal filter will take advantage of the above comment, and it will reduce strongly the noise when the blurring artifact, introduced for noise reduction, is not appreciable by the audience.

The implementation of a visual human system based filter, has a high computational cost. A modification, in order to reduce this cost, is done over the original response. In figure 1 we can see the human system response and the approximated filter response implemented.

3.2 Filter Work

In order to understand the filter performance, as follows, we will see an analysis of the filter behavior with regard to the object motions and its repercussion in the resulting quality. Seeing the spatial-temporal filter response depicted in figure 2, we are able to establish the following comments:

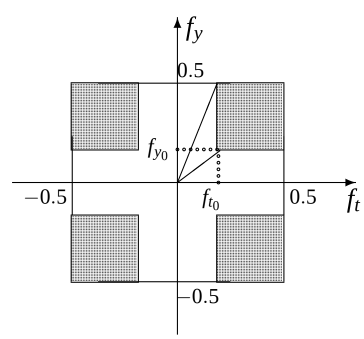


Figure 2: $f_y - f_t$ section of the spatial-temporal ideal filter response.

- Static Objects: $f_t = 0$. All of them will be inside the filter pass band. These objets remain unchanged before filtering. This performance benefits the final quality of these objects because for static objects the eye is able to detect the best detail and for these reasons the filter must not degrade them.

- Objects whose velocity is $v < \frac{f_{t0}}{0.5}$ (pixel/frame), neither suffer the blurring effect, independently of their size. f_{t0} is the filter temporal cutoff frequency.
- Objects whose velocity is $\frac{f_{t0}}{0.5} < v < \frac{f_{t0}}{f_{y0}}$ (pixel/frame). f_{y0} is the filter vertical cutoff frequency. With this velocity value, the filter begins to blur the object because the spatial bandwidth decreases when the velocity increases. However, although the object results blurred by the filter, this effect, in a motion object, can not be appreciated by the human eye.
- For velocities $v > \frac{f_{t0}}{f_{y0}}$, the blurring effect does not increase more and now, the temporal bandwidth lets start to increase. This filter parameter is imposed because objects with velocity higher than this threshold let the human eye to appreciate the temporal smoothness.

This behavior, above commented, can be also explained observing the differences between the bandwidth of a temporal filter, vs. the velocity, (depicted in figure 3) and the bandwidth of the proposed spatial-temporal filter (shown in figure 4). For low velocities, both do not filter the object. When the object motions increase (from $v < \frac{f_{t0}}{0.5}$), the spatial bandwidth decreases in both cases, reducing the noise and so, blurring the object in an inappreciable manner. However, this decrease finishes at $v < \frac{f_{t0}}{f_{y0}}$ for the spatial-temporal filter and not for the temporal one, which continues reducing the noise but also blurring the object. From this analysis we can establish that, in situations of high motion, the temporal filter reduces more noise than the spatial-temporal but, on the other hand, it also blurs more the image. The excessive blurring, is more unpleasant for the viewer than the presence of noise.

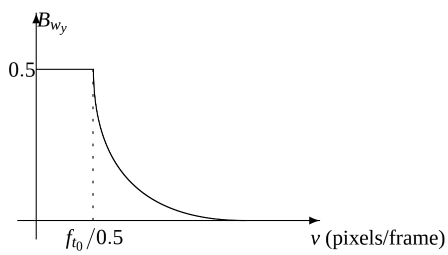


Figure 3: Spatial bandwidth vs. velocity, for the temporal filter.

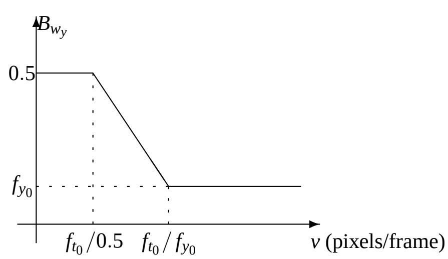


Figure 4: Spatial bandwidth vs. velocity, for the spatial-temporal filter.

3.3 Efficient Implementation

Once the filter performance has been seen, we will explain the filter implementation in order to reduce the computational cost. This reduction is based on separating the filter response with the aim of obtaining a set of separable filters.

The frequency response of the spatial-temporal filter is depicted in figure 5, where W_t, W_x and W_y represent the filter cutoff frequencies for each dimension, temporal and both spatial frequencies, respectively.

This response can also be expressed as the sum of the followings terms:

$$H(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_t}) =$$

$$H_1(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_t}) + H_2(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_t}) - H_3(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_t}) \quad (4)$$

where $H_1(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_t})$ is the frequency response of the spatial filter shown in figure 6-(a), $H_2(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_t})$ is a temporal filter whose frequency response can be observed in figure 6-(b) and, finally, the frequency response of the spatial-temporal filter $H_3(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_t})$ is depicted in figure 6-(c).

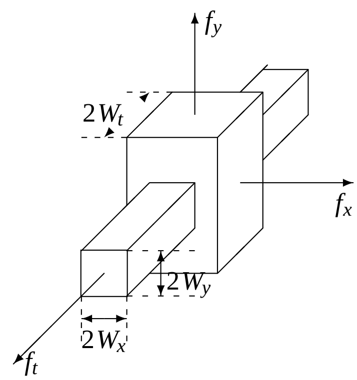


Figure 5: Frequency response of the spatial-temporal filter $H(e^{j\omega_x}, e^{j\omega_y}, e^{j\omega_t})$.

In order to implement the filtering process in the spatial-temporal domain, we have to obtain the impulse response of our spatial-temporal filter. Taking the Inverse Fourier Transform of the expression 4, the impulse response will be also obtained as sum of three terms:

$$h[x,y,t] = h_1[x,y,t] + h_2[x,y,t] - h_3[x,y,t], \quad (5)$$

where

$$h_1[x,y,t] = h_e[x,y]\delta[t]$$

$$h_2[x,y,t] = h_t[t]\delta[x,y]$$

$$h_3[x,y,t] = h_1[x,y,t] * h_2[x,y,t]$$

being

$$h_e[x,y] = \begin{cases} \frac{1}{MN}, & \text{if } |x| < \frac{M-1}{2} \text{ and } |y| < \frac{N-1}{2} \ ; \\ 0, & \text{other case.} \end{cases}$$

and

$$h_t[t] = \begin{cases} \frac{1}{L}, & \text{if } |t| < \frac{L-1}{2}; \\ 0, & \text{other case.} \end{cases}$$

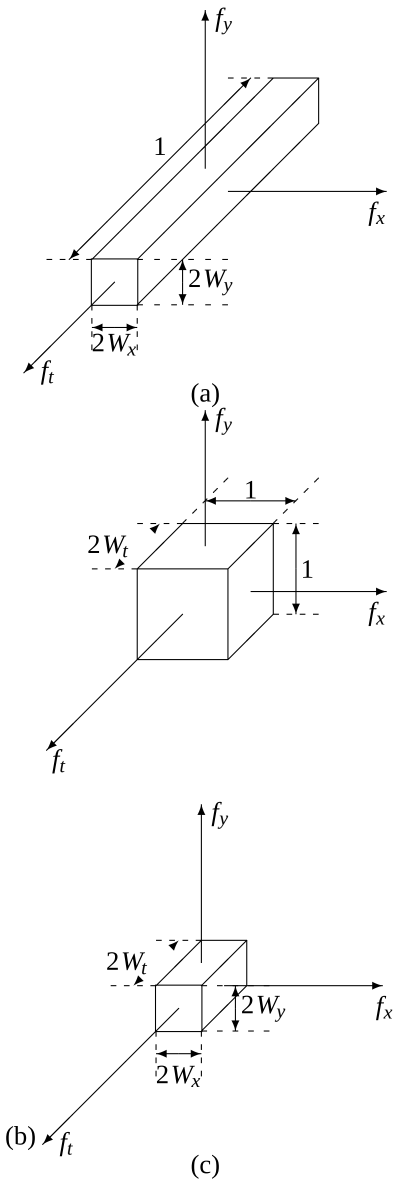


Figure 6: Frequency responses. (a) H_1. (b) H_2. (c) H_3.

As can be deduced from the previous expressions, all the different filters involved in the filtering process are separable. Attending to this property, the implementation of the proposed filter is very efficient computationally. It would be desirable to choose odd values for M, N and L in order to get zero phase filters, which do not produce image displacements.

In figure 7 we can see the real frequency response (the section $f_y - f_t$), for the implemented filter.

3.4 Filter Design Parameters

The parameters used in order to control the filter performance are the spatial size $M \times N$ and the temporal filter size L. These parameters have an influence on the filter frequency response in the following way:

- Higher values of L produce smaller f_{t0} values. For this reason, an object can have a velocity to be blurred which decreases if L arises.

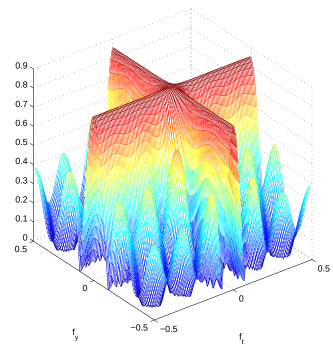


Figure 7: $f_y - f_t$ section of the real response of $FT\{h[x,y,t]\}$

- The maximum blurring effect is related with (f_{x0}, f_{y0}), and so with (M,N). If (M,N) increases (f_{x0}, f_{y0}) decreases, increasing, therefore, the maximum blurring produced.

Summarizing the points described above we can say that: changing L we are able to control the maximum object velocity to be filter without being blurred due to the temporal smoothing; and $M \times N$ let us control the maximum blurring we will have after applying the spatial-temporal filter, independently of the object velocity.

4. RESULTS

In order to quantify the improvement with the image noise level we have introduced uniform noise with several variance levels in a set of test sequences (*Foreman*, *Carphone*, *Bridge*,...). To measure the noise reduction we have used an objective parameter and another one subjective. The objective parameter used to measure the noise reduction has been the noise reduction factor (n.r.f.), which is defined as:

$$n.r.f. = \frac{\sigma_o^2}{\sigma_i^2} = \sum_x \sum_y \sum_t |h[x,y,t]|^2, \qquad (6)$$

where σ_i^2 is the input noise variance and σ_o^2 is the variance of the output noise.

Figure 8 shows the variation of the noise reduction factor, in dB, depending on the filter parameters L, N and M ($M = N$). We can observe how the noise reduction factor decreases (more noise reduction) when the filter dimensions decrease.

The table 1-(a) presents a n.r.f comparison between the spatial filter of size $M \times N$ ($h_e[x,y]$) and the proposed spatial-temporal filter using the same size for the spatial dimensions and a value of $L = 3$ for the temporal dimension. As we can observe, if $M \times N$ increases the n.r.f. decreases for both filters. However this decrease is smaller in the case of the spatial filter, so the quantity of reduced noise is higher. In the same way, the goal of the table 1-(b) is to compare the n.r.f of the temporal filter of size L ($h_t[t]$) with the spatial-temporal filter using the same size in the temporal dimension and (3×3) in the spatial dimensions. The conclusions are the same as in the case of the spatial comparison: given a certain L, the temporal filter reduces more noise than the spatial-temporal, and in both cases the noise reduction increases when L increases.

The subjective tests have been carried out using a wide audience in order to watch the set of noisy sequences (in-

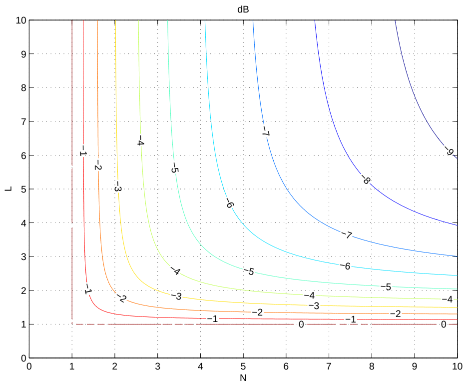


Figure 8: Noise reduction factor depending on the spatial-temporal filter size.

$M \times N$	$n.r.f_{e-t}$	$n.r.f_{\cdot e}$
3×3	-3,89	-9,54
5×5	-4,44	-13,98
7×7	-4,59	-16,9
9×9	-4,66	-19,08

(a)

L	$n.r.f_{e-t}$	$n.r.f_{\cdot t}$
3	-4,44	-4,77
5	-6,34	-6,98
7	-7,51	-8,45
9	-8,33	-9,54

(b)

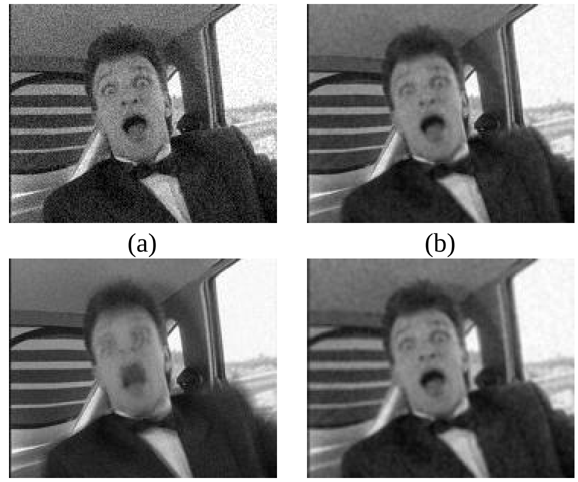
Table 1: Comparison of the Noise Reduction Factor (dB): (a) spatial-temporal filter $M \times N \times 3$ and spatial filter $M \times N$. (b) spatial-temporal filter $5 \times 5 \times L$ and temporal filter L.

troducing different noise variances) and the same sequences after being filtered using different filters (temporal, spatial and the proposed spatial-temporal) and varying the filter parameters. The information extracted from this subjective test is the kind of filter and the set of parameters for which the most of the audience prefers the sequences filtered using them. The selected filter has been the spatial-temporal with $M \times N = 3 \times 3$ and $L = 9$. In figure 9 we can observe the differences between a noise frame (from the sequence Foreman) and the results of filtering the frame using the spatial-temporal filter ($3 \times 3 \times 9$), the temporal filter ($L = 9$) and the spatial filter (3×3). We can see how when there is no motion (in the side window) the image is not blurred by the spatial-temporal filter and it gets to reduce the noise, in the same manner than the temporal filter; however the spatial filter blurs more the image. In the case of motion (the character face), the spatial-temporally filtered image suffers the same blurring than the spatially filtered image, being the temporal filter which introduces more blurring.

The set of parameters selected as the optimum has been also used in order to filter sequences from old films, with the purpose of reducing grain noise, obtaining great results.

5. CONCLUSIONS

In this paper we have presented a filter which achieves the proposed goals: to reduce noise in sequences, introducing an unavoidable degradation in the restored sequence not noticeable by a viewer. The proposed filter is a spatial-temporal FIR filter, based on the properties of the visual human system and implemented efficiently. This computational efficiency is as a result of implement the filter response as a sum of separable filters. Although the quantity of reduced noise is a little smaller than the reduction obtained using other FIR filters, only spatial or only temporal, of the same size, the degradation (the blurring) introduced in the sequence is smaller; therefore, the effect produced in a viewer is much better.



(a) (b) (c) (d)

Figure 9: Filter Comparison: (a) Original Sequence: *Carphone + Noise*. (b) Spatial-temporal filter ($3 \times 3 \times 9$). Temporal Filter ($L = 9$). (d) Spatial Filter (3×3).

REFERENCES

[1] M.K Özkan, A. Tanju Erdem, M.I Sezan, and A. Murat Tekalp. “Efficient multiframe wiener restoration of blurred and noisy image sequences.” *IEEE Transactions on Image Processing*, 1(4), pp 453-476, October 1992.

[2] Cheuk L. Chan, A. K. Katsaggelos and A. V. Sahakian.“Linear-quadratic noise-smoothing filters for quantum-limited images.” *IEEE Transactions on Image Processing*, 4(9), pp 1328-1333, Semtember 1995.

[3] J.M. Boyce. “Noise reduction of image sequences using adaptive motion compensated frame averaging, in *Procc. of the IEEE Int. Conf. Acoust., Speech, Sigal Process.*, 3 pp 461-464, San Franciasco, CA, March 1992.

[4] Richard P. Kleihorst, Reginald L. Lagendijk, and J. Biemond.“Noise Reduction of severely corrupted image sequences, in ” *Procc. of the IEEE Int. Conf. Acoust., Speech, Sigal Process.*, 5, pp 293-296, Minneapolis, Minnesota, USA; April 1993.

[5] M.B. Alp and Y.Neuvo. “3-dimensional median filters for image sequence processing, in ” *Procc. of the IEEE Int. Conf. Acoust., Speech, Sigal Process.*, 4 pp 2917-2920, Toronto, Canada, May 1991.

[6] James C. Brailean and A. K. Katsaggelos. “Recursive displacement estimation and restoration of noisy-blurred image sequences. ” *IEEE Transactions on Image Processing*, 4(9) pp 1236-1251, September 1995.

[7] Peter Van Roosmalen. “Restoration of archive films and Video. PhD Thesis, University of Delft, 1999.

[8] Eric Dubois. “The sampling and reconstruction of time-varying imagery with application in video systems. ” *Proc. of the IEEE*, 73(4) pp 502-522, April 1985.