Rapid and Reliable Detection of Film Grain Noise*

Peter Schallauer†, Roland Mörzinger

JOANNEUM RESEARCH, Institute of Information Systems and Information Management,
Steyrergasse 17, 8010 Graz, Austria

ABSTRACT

The knowledge of material specific film grain characteristics can significantly improve the performance of digital film processing algorithms. This paper proposes a rapid and reliable detector for film grain properties. First, homogeneous blocks concerning intensity and texture are determined by a novel measure calculated in the frequency domain. Second, based on these blocks the signal-dependent grain noise level is estimated and an image region of pure film grain is detected. Evaluation based on a wide range of structured images with high frequent electronic noise and low frequent film grain, shows that the method performs with a worst case estimation error of 2.54 dB in typical movie quality (PSNR of 25-40 dB). This meets the requirements of real applications such as digital film restoration and special effects compositing.

Index Terms— random noise, signal dependency, homogeneous textures, digital film restoration

1. INTRODUCTION AND RELATED WORK

Film grain is an inherent artifact of analog film stock and has a significant impact on the effectiveness of digital film processing. When information about film grain is available, the quality of film reconstruction, restoration, compositing and compression for broadcast delivery and DVD production can be improved significantly [1]. Film grain characteristics of interest are the amount of grain, the signal dependency of grain and the detection of grain regions containing pure film grain. These *grain template* regions are a pre-requisite for high quality film grain synthesis [1].

Film grain is linked to the physical character of photographic film and is perceived as a random pattern based on local-density variations in an area of uniform density, cf. Figure 4. These density differences result from the random grouping of silver particles into denser and less dense areas in the film's emulsion layer [2].

Noise can be modeled as a stochastical signal (Gaussian or non-Gaussian) which is additive or multiplicative to an image signal. It can further be modeled signal-dependent or signal-independent. The most advanced film grain noise models proposed in [3,4,5] assume that signal dependency can be modeled in a way that the amount of grain noise is proportional to a power p of the optical density with $0 \leq p \leq 1$. This fixed proportionality is a strong limitation considering that signal dependency depends on the film stock, the film lighting conditions, the non-linear relation between film exposure and density and on the non-linear operations possibly applied within the digitization process, e.g. color correction. For this reason we propose modeling signal dependency by a continuous polynomial function, which overcomes this strong restriction. For measuring noise, different block-based approaches have been suggested. There, noise characteristics are estimated from a set of homogeneous blocks (e.g. with low variance) in the image, e.g. [6] presents an adaptive block-based approach for estimating the video noise level. However, this method would consider film grain as image structure (edges, corners), because film grain has a larger spatial dimension than electronic noise. We suggest a block-based technique, considering specific features of film grain noise. Moreover our approach aims at detecting a spatial grain template region and information about the signal dependency of film grain.

2. DETECTION OF FILM GRAIN NOISE

The proposed grain detection algorithm assumes that (1) the material specific appearance of film grain can most clearly be captured in homogeneous areas of the material, (2) preferably in the mid-gray areas of the film, (3) film grain is spatially global, (4) film grain is a texture with similar characteristics in all directions and has therefore no dominant direction, (5) in the frequency domain, film grain is represented in mid-high to highest frequencies, (6) the variation in image intensities is a measure for the grain level and (7) contrary to electronic noise, the spatial dimension of film grain can be significantly larger than one pixel depending on sampling resolution.

The proposed grain detection framework consists of a homogeneity analyzer, a signal dependency estimator and a grain template detector. The homogeneity analyzer determines the homogeneity of a block in terms of intensity and texture. The signal dependency estimator yields a function of grain level over signal intensity. The grain template detector extracts a grain region, which is an area without image structures containing only pure film grain.

A. Block Homogeneity Analysis

Homogeneity in a square block B of size w x w is analyzed in regard to image intensities and directionality.

In order to take local variations within B into account, the intensity homogeneity measure H_{intB} is defined as the maximum standard deviation σ from either the whole block B or its non overlapping quarters B_i of size $w/2$ x $w/2$:

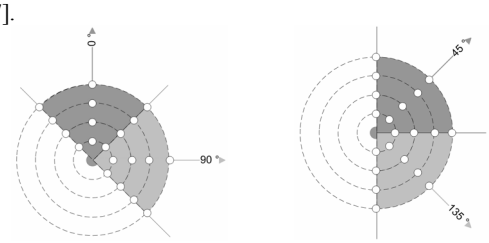
$$H_{\mathrm{int}\,B} = \max\begin{pmatrix} \max\left(\sigma_{B_i}\right) \\ \sigma_B \end{pmatrix}, \; i = 1,\ldots,4$$

For estimating the directionality of a block's intensity pattern, we analyze the low frequencies in B that are not due to film grain.

Direction information of a block's content is preserved in its frequency domain representation, e.g. its 2D Fourier spectrum.

† peter.schallauer@joanneum.at; phone +43 316 876-1202; fax +43 316 876-1191; www.joanneum.at/iis/en

Note that we discard the phase information and use the Fourier *log* power spectrum without normalization, cf. Figure 2. The frequency domain representation can be scanned by a revolving wedge pointed in the zero frequency component, spanning a radial band[7].



a) b)

Figure 1 Directionality calculation map centered at Fourier power spectrum. The shown four shaded angular wedges each span 90° in the directions of (a) 0° and 90° and (b) 45° and 135°. The black and white spots (12 per wedge) denote locations for frequency magnitude values. The dashed lines indicate the different radii (here: 1.5, 2.5, 3.5, 4.5 units distance from center).

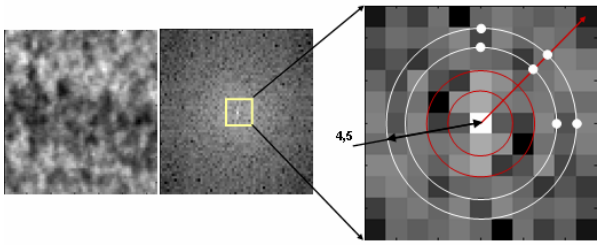


Figure 2 (Left) Square block, (middle) its 2D Fourier log power spectrum, (right) directionality calculation map applied on the 11x11 pixels central section. For the measure $\theta_{[3.5:4.5]45°}$, the locations of the frequency magnitude values are marked.

The radial band is described as $r = [l:h]$, where l is the lowest frequency component and h the highest frequency component considered. A band $r = [1:5]$ captures the orientation of the visually most relevant image components [8]. It is justified to fix the values of l and h provided that w is adapted to the material's resolution so that film grain is mapped to the parts of the spectrum beyond h. We have concluded on a large set of grain templates that a block size w assigned to $y/15$, where y is the source image height, is required to sufficiently capture the lowest frequency components of film grain.

We propose overlapping wedges, each of which explore a 90° angle directed at α = {0°, 45°, 90°, 135°}. Full orientation is covered owing to the symmetric property of the power spectrum. A wedge directed to α contributes to the measure θ_{ra}, where r is either assigned to the radial band of $r=[1.5:2.5]$ (considering the lowest frequencies around zero frequency) or $r=[3.5:4.5]$ (for the next higher frequencies). The measure θ_{ra} is a weighted mean of six representative frequency magnitude values (cf. black or white spots in Figure 1), where the four values at the wedge's border are weighted half as much (0.125 each) as the two inboard ones (0.25 each). The center of a frequency magnitude value may not match a pixel's center in the power spectrum. Therefore, bicubic interpolation is done. Figure 1 shows a map with the 4 different radii plotted.

The homogeneity criterion in regard to directionality H_{dirB}^{r} for a certain radial band r is defined as

$$H_{dirB}^{r} = \max(\theta_{r\alpha} - \mu_r) - \max(\min(\theta_{r\alpha} - \mu_r), 0), \quad \alpha = 0°, 45°, 90°, 135°,$$

where θ_{ra} is the directionality measure for direction α and μ_r denotes the average of all frequency magnitude values in r.

B. Template and Signal Dependency Detection

Detection starts with randomly sampling positions for blocks B. A *pixel skipping* factor reduces the number of possible window locations by creating a grid space to which each block is aligned. The homogeneity analyzer presented in Section *A* gives measure about the homogeneity of a block B (H_{intB} and H_{dirB}^{r}). Any B whose measures fall below the initial rough thresholds tH_{intB} and tH_{dirB}^{r} is added to a set of so-called *grain blocks*. The global threshold tH_{intB} is not crucial, merely facilitates pre-selection and is refined in the further process. For tH_{intB} an initial value of 15 has proved to yield good results, and for all tH_{dirB}^{r} we empirically determined the thresholds to be 0.4. Note that throughout the paper, 8bit quantization of the material is assumed.

The intensity values range (0 to 255) is divided into k bins of equal size. Any block B is assigned to exactly one of the k bins, based on its mean intensity value μ_B. The threshold tH_{intB} is adapted for each of the k bins. The new intensity threshold for a bin is assigned to 120% of the minimum H_{intB} (cf. assumption 3 in Section 2) of all the grain blocks that have been assigned to the bin. Subsequently region growing is performed. It benefits from the adaptive signal-dependent threshold and can therefore be responsive to material specific properties that are not a priori known, e.g. the overall amount of film grain. The region growing dilates all grain blocks within the limits of the grid space in a 4-neighbourhood, resulting in a set of connected and overlapping grain blocks, called *grain regions*. The dilation process produces several interim grain regions, each of which is evaluated as described below.

To estimate the probability that a grain region is composed only and entirely of film grain, the following rating concept for a single block or a set of blocks is proposed. The rating q_V for a set of blocks V, containing one or more connected and overlapping grain blocks is defined as

$$q_V = \frac{1}{N} \sum_{B \in V} \frac{1}{4} \sum_{i=1}^{4} 1 - \frac{|tH_i - H_{Bi}|}{tH_i},$$

where N is the number of grain blocks in V, H_B is $H_B=\{H_{intB}, H_{dirB}^{[1.5:2.5]}, H_{dirB}^{[3.5:4.5]}, \mu_B\}$ an array containing measures of B and tH is $tH=\{ tH_{intB}, tH_{dirB}^{[1.5:2.5]}, tH_{dirB}^{[1.5:2.5]}, \mu_d\}$ an array with the thresholds and the desired mean grain template intensity μ_d, and i is a subscript identifying elements in the arrays H_B or tH. This rating expresses the probability by a mean normalized difference between the four measures and their thresholds on a scale from 0 to 1.

The final grain template region is the largest inscribed rectangle computed in the highest rated grain region.

In the course of signal dependency estimation, for each of the k intensity range bins, the minimum detected grain level (H_{intB}) is computed from the grain blocks of all grain regions assigned to the certain bin, resulting in k discrete grain levels at most. Further, a polynomial of 4th degree is estimated that best fits these grain levels. The computation of the polynomial in a least mean square sense assumes additional discrete grain levels of value 0.5 for the intensities 15 and 240.

3. EVALUATION

We have used the following methodology to evaluate our algorithm, which is similar to common evaluation procedures for noise estimation.

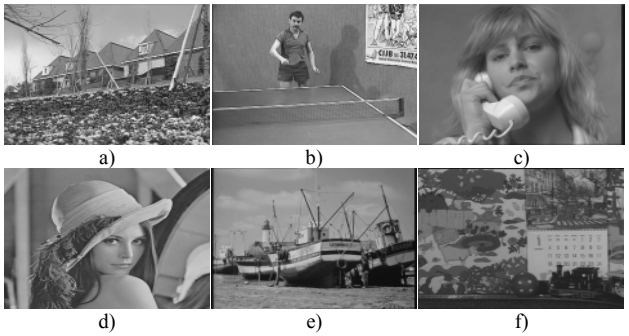


Figure 3: Test images/sequences referred to in the evaluation: a) garden b) tennis c) susie d) lenna e) boats f) mobile.

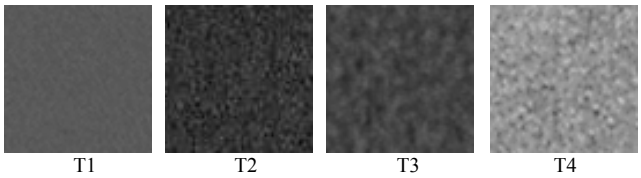


Figure 4: Templates for film grain. Note that different granularities from very fine to coarse are supported.

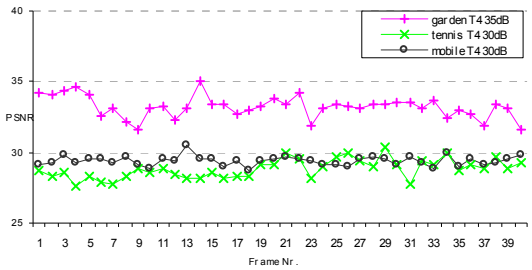


Figure 5: Single frame PSNR over time for different sequences. Different grain types (T3, T4 from Figure 4) and levels (35 dB, 30 dB) are applied.

	μ PSNR	σ PSNR	Max. deviation (dB)
garden 35dB T3	33.21	0.78	3.44
tennis 30dB T3	28.83	0.69	2.45
mobile 30dB T4	29.39	0.34	1.26

Table 1: Statistical variability of grain template PSNR over time.

Firstly, film grain is synthesized based on grain templates [1] shown in Figure 4, and superimposed producing a specified Peak-Signal-Noise-Ratio (PSNR). For higher precision, the material is cleaned up from prior grain and noise by median filtering. Secondly, grain properties are detected, and thirdly, the differences between the known properties and the detected ones are analyzed. Superimposition is done by pixelwise adding the synthesized grain to the image content. For evaluating the signal dependency estimation, synthesized grain is added according to a given signal dependency function, which is based on the intensity of the 9-tap median filtered image content.

The test material in this paper shown in Figure 3 is available to the public. For simulating real-world application conditions, particularly for runtime measurement, the original images and frames are enlarged prior to superimposition, resulting in a resolution of 1024x1024 pixels for *lenna*, 1440x1152 for *boat*, 720x576 for *mobile* and *susie*, 704x480 for *garden* and *tennis*.

Figure 5 demonstrates the stability of grain level detection in an image sequence over time.

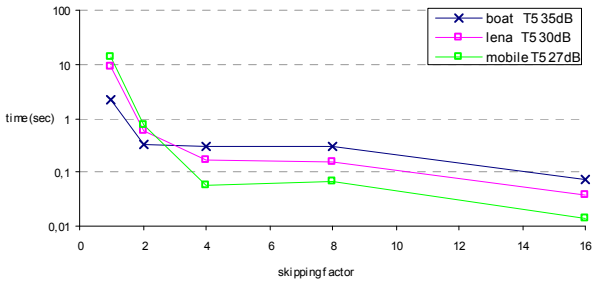


Figure 6: Speed optimization test. Reduction of execution time due to increased pixel skipping factor. Shown time measurements are averaged over 10 executions per test image. Content features grain type T4 from Figure 4 with different PSNR (27, 30, 35 dB).

	μ PSNR	σ PSNR	Max. deviation (dB)
Boat 35dB	34.21	0.24	1.14
Lenna 30dB	29.91	0.20	0.40
Mobile 27dB	26.99	0.12	0.18

Table 2: Statistical variability of speed optimization test.

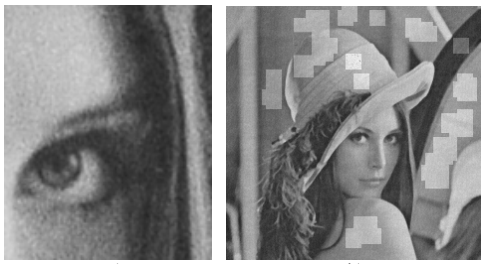


Figure 7: Signal dependency test. a) Detail of lenna after 9-tap median filtering and subsequent adding of grain of type T2, b) detected grain blocks used in signal dependency estimation.

It shows the PSNR of grain templates, which have been detected for each frame in different image sequences, exhibiting varying content and different types and amount of film grain. Generally, the grain level is overestimated, i.e. average *μPSNR* is estimated lower than the known PSNR of synthesized grain. Overestimation is more apparent in *garden*, and less in the cleaner *mobile* sequence which indicates that this effect is mainly caused by the superimposition of existing noise and synthesized grain. Table 1 shows the test's variability; the maximum standard deviation of PSNR is 0.78 and the maximum single frame PSNR error (deviation from input PSNR) is 3.44 dB.

Figure 6 shows the computational time as a function of the pixel skipping factor (c.f. Section 2.B). The three test images have been subjected to a 9-tap median filter prior to grain detection. The differences in execution time are due to the number of homogeneity analyses (see Section 3.A), which vary from approx. 14.000 in the slowest case with skipping factor 1, to 400 in the fastest case with factor 16. Table 2 shows that the results of grain level detection remain stable and accurate despite of a larger pixel skipping factor.

For the evaluation of signal dependency detection, the *lenna* image is used, as it contains a mixture of detail, flat regions, shading and

texture. Synthesized grain is added according to a given input signal dependency considering typical film stock characteristics, e.g. an almost zero grain level in dark and bright image areas and a Gaussian like grain level distribution in the mid tones.

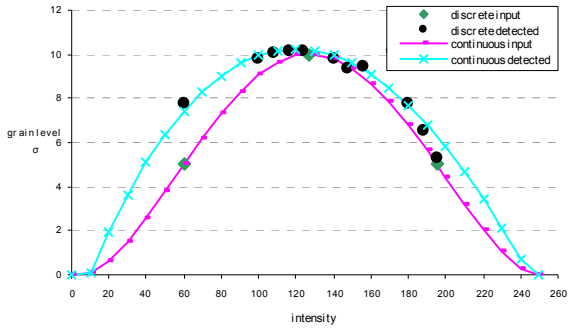


Figure 8: Signal dependency detection on *lenna* image.

$PSNR_o$	σ_o	$\mu_{E\sigma}$	$\sigma_{E\sigma}$	μ_{Epsnr} (dB)
40	2.50	0.91	0.25	2.54
37.5	3.32	0.68	0.20	1.41
35	4.43	0.79	0.37	1.22
32.5	5.91	0.76	0.44	0.85
30	7.84	0.73	0.26	0.53
27.5	10.45	0.87	0.22	0.45
25	14.01	0.43	0.20	0.06

Table 3: Statistical variability of the estimation error as a function of film grain standard deviation σ_o, or $PSNR_o$ respectively. μ_{Epsnr} is the mean error between known and detected PSNR. Given values are averaged over 5 runs.

The effect of signal dependency is visualized in Figure 7; it shows both, a detail of *lenna* exhibiting the graininess synthesized by template T2, and the grain block regions that have been detected and used for the signal dependency estimation. Note that less graininess is perceived in the very bright and dark image details, whereas grain is pronounced in the mid gray areas (e.g. below *lenna*'s eye).

Figure 8 shows the input signal dependency, conveyed by three discrete grain levels and its derived continuous signal dependency representation (for information about its computation see end of Section 2.0). Further, it shows the detected grain levels and the resulting continuous signal dependency. The mean error of signal dependency estimation is 1.15, which is equivalent to a PSNR error of 3.14 dB. The figure shows that this test achieves a good estimate of the maximum grain level in the mid tones. In the *lenna* image it is more difficult to detect signal dependency accurately in darker and brighter areas. This capability highly depends on the spatial distribution of intensities within the image content.

For evaluating the overall error in grain level detection, six still images are used, shown in Figure 3. After 9-tap median filtering, the images are superimposed by synthesized film grain with a PSNR typical for film grain, that is, 25 to 40 dB (corresponds to adding a standard deviation of film grain noise σ_o from 2.5 to 14.01). The estimation error is defined as the absolute difference between known and detected standard deviation: $E_\sigma=|\sigma_o - \sigma_D|$. The results of the test are summarized in Table 3. The results show that the proposed method performs with low PSNR estimation errors (maximum $\mu_{E\sigma}$=0.91 or μ_{Epsnr}=2.54 dB respectively) for both high and low grain levels.

4. CONCLUSION AND APPLICATION

Film grain has other specific characteristics than electronic noise: its magnitude is signal-dependent and the spatial expansion is typically wider. These characteristics are being considered with our method of detecting properties of film grain and the grain template. The proposed film grain detection algorithm is suited for the estimation of high frequent electronic noise, whereas generally electronic noise detectors are not designed for detecting film grain properties.

The proposed algorithm shows a worst-case estimation error of 2.54 dB and allows stable grain level detection over time. Owing to pixel skipping, the detection is executed on HD resolution within a matter of a few hundred milliseconds on an average, without degrading the quality of detection. The reliability of signal dependency estimation is subject to the availability of homogeneous regions at various intensities within the image content. Future efforts need to be made to detect grain properties also in structured image content.

Small detection errors enable practical applications for still images as well as film and video material. The detection and matching of grain properties is important in post production where perceptually similar grain is added to computer generated content during compositing, in digital film reconstruction for matching grain from different copies and in digital film restoration for adapting grain of defect removed areas. The proposed approach is based on grain templates and thus enables a lossless transfer of film grain information between the steps of detection and matching. Further applications can be seen in film and video quality assessment, e.g. as a pre-requisite for efficient grain suppression for DVD production and video broadcast delivery, c.f. MPEG-4/H.264 SEI messages for film grain encoding. Future work may be needed for improving detection reliability in structured image content. For that purpose, use of spatio-temporal, inter-frame information can be made.

5. ACKNOWLEDGEMENTS

The authors would like to thank Werner Bailer, Georg Thallinger and Werner Haas as well as several other colleagues at JOANNEUM RESEARCH, who contributed valuable feedback to the work. This work has been funded partially under the 6th Framework Programme of the European Union within the IST project "PrestoSpace" (IST FP6-507336, http://www.prestospace.org).

6. REFERENCES

[1] P. Schallauer, R. Mörzinger, "Film Grain Synthesis and its Application to Re-graining," *Proc. Conf. Image Quality and System Performance III, Electronic Imaging*, January 2006.

[2] Kodak. "Kodak Motion Picture Film: H-1," *Eastman Kodak Company*.

[3] S. Chen, J. Murray, J. Walkup, "Robust image estimation in signal-dependent noise," *Proc. ICASSP 84*, pp. 585-588, March 1984.

[4] J. Yan, D. Hatzinakos, "Signal-dependent film grain noise removal and generation based on higher-order statistics," *IEEE Signal Processing Workshop on Higher-Order Statistics*, pp. 77-81, July 1997.

[5] B. Aiazzi, L. Alparone, S. Baronti, "A robust method for parameter estimation of signal-dependent noise models in digital images," *Proc. DSP97*, pp. 601-604, July 1997.

[6] A. Amer, E. Dubois, "Fast and Reliable Structure-Oriented Video Noise Estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 113-118, January 2005.

[7] K. Lee, L. Chen, "An efficient computation method for the texture browsing descriptor of MPEG-7," *International Journal of Computer Vision, Vol.23, No.5,* pp. 479-489, May 2005.
[8] A. Celentano, V. Lecce, "A FFT based technique for image signature generation," *Proc. Storage and Retrieval for Image and Video Database, SPIE Vol.3022*, February 1997.